Reducing Metal Streak Artifacts in CT Images via Deep Learning: Pilot Results

Lars Gjesteby, Qingsong Yang, Yan Xi, Bernhard Claus, Yannan Jin, Bruno De Man, Ge Wang*

***Abstract*—Machine learning including deep learning is rapidly gaining popularity as a generic solution to problems across many fields. In medical imaging, deep learning has been successfully applied to image processing and analysis. In this paper, we employ a convolutional neural network (CNN) to address the long-standing problem of metal artifacts in CT images. Despite a vast number of metal artifact reduction (MAR) methods developed over the past four decades, there remain clinical areas in need of better results. Specifically, proton therapy planning requires high image quality for accurate tumor volume estimation. Errors in the image reconstruction may lead to treatment failure. In this paper, we merge deep learning with a state-of-the-art normalization-based MAR algorithm, NMAR, to correct metal streaks in critical image regions. Our results show that deep learning is a novel way to address CT reconstruction challenges, yielding images superior to the state of the art.**

***Keywords*—Computed tomography (CT), deep learning, convolutional neural network (CNN), metal artifact reduction (MAR), proton therapy planning.**

I. INTRODUCTION

Metal artifacts are a long-standing problem in CT that severely degrade image quality. Many MAR techniques were developed over the past four decades (see [1] for a comprehensive overview), but their translation to clinical settings is not always feasible or totally successful. For those algorithms that have been adopted clinically, there remain important applications in which a sufficient image quality cannot be achieved, such as for proton therapy planning. Tumor volume estimation is very sensitive to image reconstruction errors, and miscalculation due to metal artifacts may result in either tumor recurrence or radiation toxicity [2], [3].

A most widely developed class of MAR methods is projection completion, in which corrupted data inside the metal trace is refined in the sinogram domain. The new data is often synthesized by an interpolation technique [4]–[7], reprojection from a prior image [8]–[11], or a combination of both that involves normalization [12]–[14]. Among these, NMAR is considered a state-of-the-art method that employs interpolation and normalization to correct data in the metal trace [13].

Other types of MAR techniques include data acquisition improvement, physics-based pre-processing, iterative reconstruction, and image post-processing. While image post-processing algorithms have had only limited success [15], [16], their merits are better seen when combined with projection-domain correction [17]. None of these methods are satisfactory for challenging applications, especially proton therapy planning.

Deep learning is a novel approach for reducing metal artifacts in CT images. The field of deep learning has risen rapidly in recent years to perform many complicated tasks with new twists [18]. Deep networks, such as a CNN, are powerful in their ability to extract detailed features from large datasets, enabling great successes in image processing and analysis [19]–[21]. In a supervised learning process, the network is trained with labeled data/images to learn how to map features between the input and the label. Once trained, the network uses forward prediction to estimate an output given an unlabeled input.

The main objective of our project is to reduce streak artifacts in critical image regions outside the metal object by combining a CNN with the state-of-the-art NMAR method. We train the network to create an end-to-end mapping of patches from metal-corrupted CT images to their corresponding artifact-free ground truth. Since raw projection data is not always accessible in commercial scanners, these initial experiments are done via numerical simulation to demonstrate the feasibility and merits of deep learning for MAR.

II. METHODS

A. Data generation

All training and test data for the CNN was generated using an industrial CT simulation software, CatSim (*General Electric Global Research Center, Niskayuna, NY*) [22]. Simple hip phantoms were defined with elliptical areas of water and bone in random aspect ratios. A metal ellipse of either titanium or stainless steel was set near each bone region to represent the hip implants. Two scans were simulated for each phantom based on a GE LightSpeed VCT system architecture. The first scan followed a standard clinical protocol, with key parameters including a tube voltage of 120 kVp, a tube current of 300 mA, 10^8 photons, 300 detectors, and 360 views at uniform projection

angles between 0-360 degrees. The 512x512 reconstructed image contained severe artifacts. For the initial correction, the image was reconstructed using the NMAR algorithm. This NMAR result served as the input to the network. A second scan on each phantom was simulated with the same parameters, except that a sufficiently higher number of photons were assigned a single energy of 70 keV to generate a monoenergetic image without artifacts. This served as the ground truth and target of the network. A total of 5,000 phantoms containing titanium or stainless steel implants were scanned. Fig. 1 displays an example image of each case with the streak regions outlined by dotted boxes. From the full images, two million patches of size 32x32 were extracted from the streak regions to form the dataset for training the CNN.

B. Network design and training

The network structure used in this paper was inspired by the study on image super-resolution in which a low-resolution input image was mapped to a high-resolution output [23]. It is a simple CNN with six convolutional layers. The first five layers contain 32 filters and a 3x3 kernel to extract features from the input patch and map them to the target patch. Each of the layers is followed by a rectified linear unit (ReLU) to introduce non-linearity. The last layer sums the estimates with 1 filter and a 3x3 kernel to form the final output patch of size 20x20. Fig. 2 illustrates the network structure.

Training was completed in the *Caffe* framework [24]. The base learning rate was 10^{-4} with a reduction factor of 0.5 after every 50k iterations. One million training iterations were carried out to minimize the loss, which is the mean squared error between the input and the target patch for each training iteration (batch size of 50).

III. RESULTS

The network performance was tested using patches from the streak regions (dotted boxes) in Fig. 1. The NMAR corrected images served as the input to the network, and the forward prediction process mapped the input to the ground truth. The ground truth and input streak regions extracted from Fig. 1 are shown in Fig. 3 for the representative titanium and stainless steel cases. The CNN prediction significantly reduced streaking artifacts for both types of metal implants.

It can be observed that the titanium case was mapped well to the ground truth, given that the input contained much less severe streaks than the stainless steel case. For stainless steel, although artifacts remain, the CNN still recovered much of the underlying background. Quantitative analysis shows that the CNN yielded substantially higher image quality than NMAR alone as measured by structural similarity index (SSIM) and peak signal-to-noise ratios (PSNR).

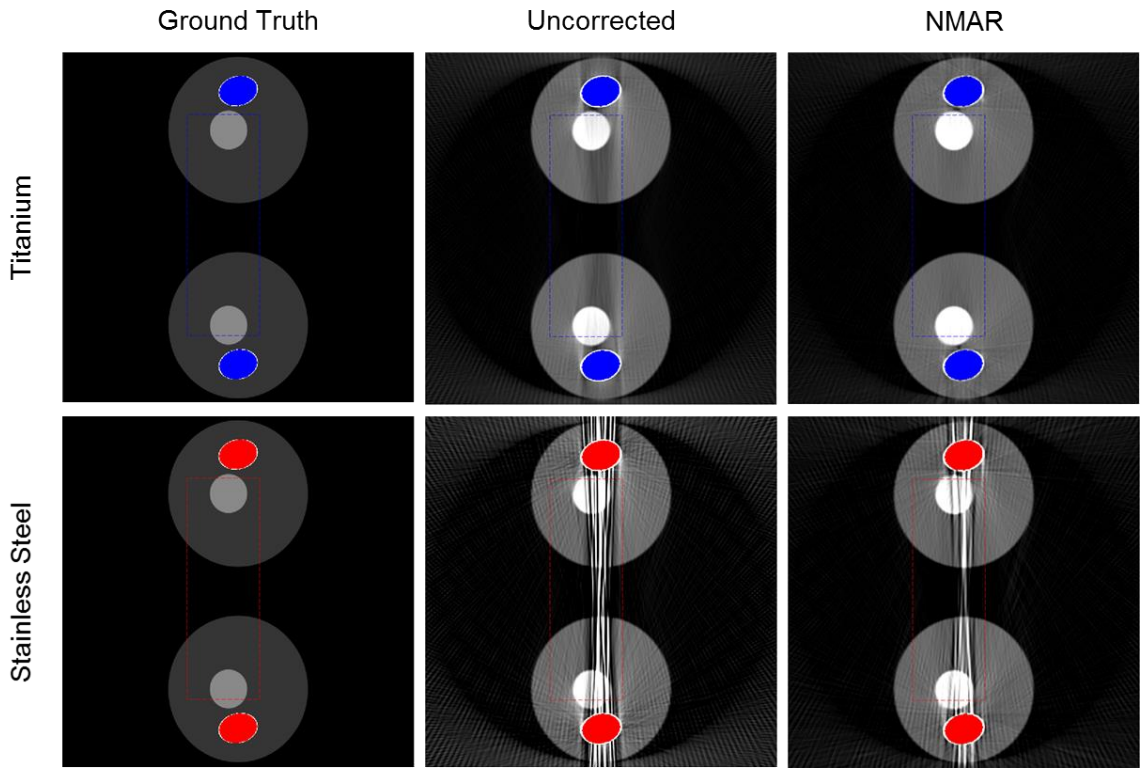

Fig. 1. Test images generated by CatSim in the cases of the titanium (blue) and stainless steel (red) implants respectively. The ground truth is a 70 keV monoenergetic image (as the network target), the uncorrected image is the initial reconstruction from a 120 kVp scan, and the NMAR image is the corrected reconstruction using the NMAR algorithm (as the network input). The dotted boxes outline the streak regions from which patches were extracted for network training.

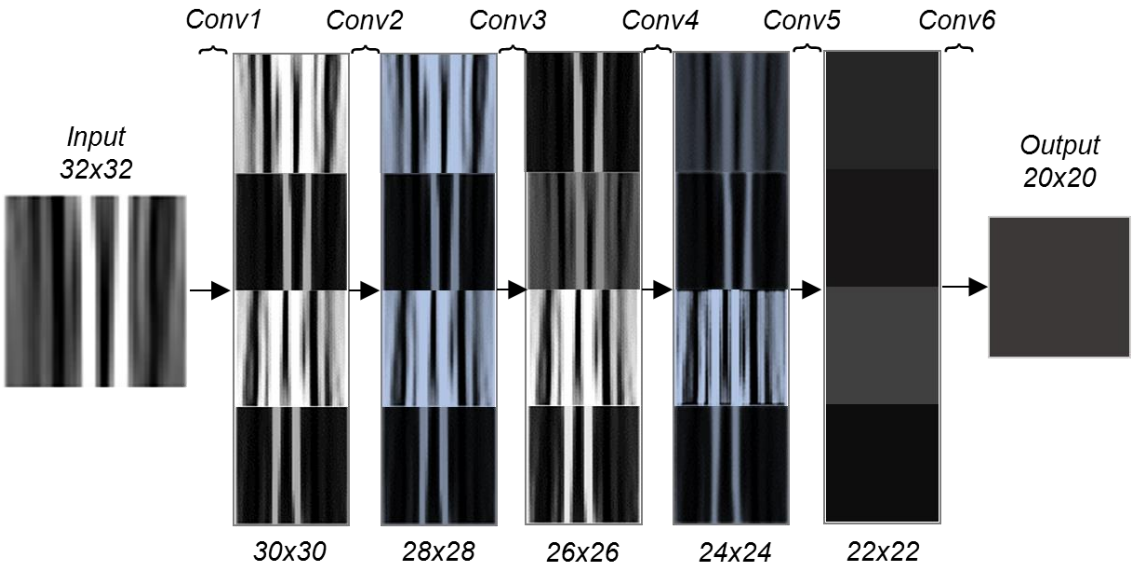

Fig. 2. Convolutional neural network containing six convolutional layers. Layers 1 through 5 have 32 filters and a 3x3 kernel, while the sixth layer has 1 filter and a 3x3 kernel. The first five layers are followed by a rectified linear unit for non-linearity. Features are extracted from a 32x32 input patch. The output prediction from these features gives a 20x20 patch.

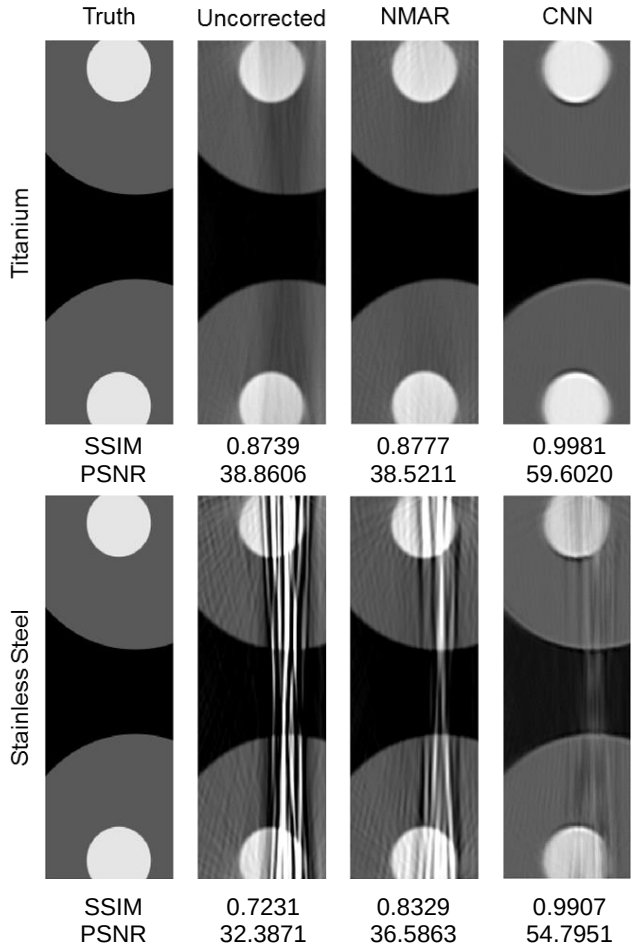

Fig. 3. CNN results for streak regions from Figure 1 and corresponding image quality metrics in reference to the ground truth. The NMAR image serves as the network input so that the MAR results will be better than the state of the art.

IV. DISCUSSIONS AND CONCLUSION

This pilot study has demonstrated that deep learning is a novel way to correct metal artifacts in CT images, and this improvement is in addition to the state of the art performance achieved using any existing approach. Our CNN combines with the state-of-the-art NMAR algorithm to provide substantial reduction of streaks in critical image regions. Visually, the best results were achieved for titanium implant cases. Titanium produces less severe artifacts than stainless steel [25], and the network has an easier path to achieving effective correction. In the case of stainless steel implants, the streaks were more severe, and the CNN prediction still corrected for residual artifacts but less effectively. Our results are promising in that there is a hope for further recovery of corrupted data beyond what NMAR achieved alone.

There is still great room for improvement in the MAR deep learning process, particularly for the challenging implant types. Our results may be improved by increasing the size and complexity of the training dataset and modifying the network structure to better train on residual error functions [26]. With more efforts, we will test images with more realistic tissue background and complex implant structures.

In summary, our work has indicated that deep learning can have an important role in achieving better image quality for MAR and other specific CT tasks, and in particular may enable more accurate tumor volume estimation for proton therapy planning.

REFERENCES

[1] L. Gjesteby, B. De Man, Y. Jin, H. Paganetti, J. Verburg, D. Giantsoudi, and G. Wang, "Metal Artifact Reduction in CT: Where Are We After Four Decades?," *IEEE Access*, vol. 4, pp. 5826–5849, 2016.

[2] G. X. Ding and W. Y. Christine, "A study on beams passing through hip prosthesis for pelvic radiation treatment," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 51, no. 4, pp. 1167–1175, 2001.

[3] C. Reft, R. Alecu, I. J. Das, B. J. Gerbi, P. Keall, E. Lief, B. J. Mijnheer, N. Papanikolaou, C. Sibata, and J. Van Dyk, "Dosimetric considerations for patients with HIP prostheses undergoing pelvic irradiation. Report of the AAPM Radiation Therapy Committee Task Group 63," *Med. Phys.*, vol. 30, no. 6, pp. 1162–1182, 2003.

[4] R. M. Lewitt and R. H. T. Bates, "Image-reconstruction from projections. III. Projection completion methods (theory)," *Optik (Stuttg).*, vol. 50, no. 3, pp. 189–204, 1978.

[5] T. Hinderling, P. Ruegsegger, M. Anliker, and C. Dietschi, "Computed Tomography Reconstruction from Hollow Projections: An Application to In Vivo Evaluation of Artificial Hip Joints," *J. Comput. Assist. Tomogr.*, vol. 3, no. 1, pp. 52–57, 1979.

[6] G. H. Glover and N. J. Pelc, "An algorithm for the reduction of metal clip artifacts in CT reconstructions.," *Med. Phys.*, vol. 8, no. 6, pp. 799–807, 1981.

[7] W. A. Kalender, R. Hebel, and J. Ebersberger, "Reduction of CT artifacts caused by metallic implants," *Radiology*, vol. 164, no. 2, pp. 576–577, 1987.

[8] C. R. Crawford, J. G. Colsher, N. J. Pelc, and A. H. R. Lonn, "High speed reprojection and its applications," in *SPIE Medical Imaging II*, 1988, pp. 311–318.

[9] R. Naidu, I. Bechwati, S. Karimi, S. Simanovsky, and C. Crawford, "Method of and system for reducing metal artifacts in images generated by x-ray scanning devices," 6,721,387 B1, 2004.

[10] C. S. Olive, M. R. Klaus, V. Pekar, K. Eck, and L. Spies, "Segmentation aided adaptive filtering for metal artifact reduction in radio-therapeutic CT images," in *SPIE Medical Imaging 2004*, 2004, vol. 5370, pp. 1991–2002.

[11] C. Lemmens, D. Faul, and J. Nuyts, "Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion," *IEEE Trans. Med. Imaging*, vol. 28, no. 2, pp. 250–60, 2009.

[12] J. Müller and T. M. Buzug, "Spurious structures created by interpolation-based CT metal artifact reduction," in *SPIE Medical Imaging*, 2009, p. 72581Y–72581Y.

[13] E. Meyer, F. Bergner, R. Raupach, T. Flohr, and M. Kachelrieß, "Normalized metal artifact reduction (NMAR) in computed tomography," *Med. Phys.*, vol. 37, no. 10, pp. 5482–5493, 2010.

[14] E. Meyer, R. Raupach, B. Schmidt, A. H. Mahnken, and M. Kachelriess, "Adaptive Normalized Metal Artifact Reduction (ANMAR) in Computed Tomography," in *2011 IEEE Nuclear Science Symposium Conference Record*, 2011, pp. 2560–2565.

[15] G. Henrich, "A simple computational method for reducing streak artifacts in CT images," *Comput. Tomogr.*, vol. 4, no. 1, pp. 67–71, 1980.

[16] M. Bal, H. Celik, K. Subramanyan, K. Eck, and L. Spies, "A radial adaptive filter for metal artifact reduction," in *SPIE Medical Imaging 2005*, 2005, vol. 5747, pp. 2075–2082.

[17] O. Watzke and W. A. Kalender, "A pragmatic approach to metal artifact reduction in CT: Merging of metal artifact reduced images," *Eur. Radiol.*, vol. 14, no. 5, pp. 849–856, 2004.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[19] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[21] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.

[22] B. De Man, S. Basu, N. Chandra, B. Dunham, P. Edic, M. Iatrou, S. McOlash, P. Sainath, C. Shaughnessy, and B. Tower, "CatSim: a new computer assisted tomography simulation environment," in *Medical Imaging*, 2007, p. 65102G–65102G.

[23] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016.

[24] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 675–678.

[25] N. Haramati, R. B. Staron, K. Mazel-Sperling, K. Freeman, E. L. Nickoloff, C. Barax, and F. Feldman, "CT scans through metal scanning technique versus hardware composition," *Comput. Med. Imaging Graph.*, vol. 18, no. 6, pp. 429–434, 1994.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.