Metal-Artifact Reduction Using Deep-Learning Based Sinogram Completion: Initial Results

Bernhard E. H. Claus, Yannan Jin, Lars A. Gjesteby, Ge Wang, Bruno De Man

***Abstract*—Metal artifact reduction (MAR) remains a hard problem with remaining limitations in image quality after more than three decades of study. Recent successes in deep learning achieve state-of-the-art performance for a range of extremely difficult and complex problems. In this work an application of deep learning to the problem of metal artifact reduction via estimation of missing data in the sinogram domain is investigated. Initial results obtained with projection data simulated from simple geometric objects show promising improvements in the quality of the estimated data in the sinogram, which is visually clearly superior to a linear interpolation approach. This improvement also translates into reduced streaking and banding artifacts in the corresponding reconstructed images. These encouraging results clearly demonstrate the potential of deep learning for addressing MAR and similar missing data problems.**

***Index Terms*—CT, Computed Tomography, Deep Learning, Metal Artifacts, MAR, Missing Data Estimation, Sinogram Completion, Sinogram Interpolation.**

I. Introduction

The field of Deep Learning is currently undergoing rapid growth, with state-of-the-art performances achieved in a range of fields, e.g., for image recognition [1], speech recognition [2], and other applications (see, e.g., [3] for an overview). Many recent advances in deep learning are addressing "big data"-type problems. Deep learning networks can -seemingly without careful engineering and domain expertise required by their designers- extract and model features of higher and higher complexity as the data propagates through the layers of the network. Due to this ability to discover intricate structures in large sets of high-dimensional data, deep learning has (potential) applicability in many domains in science and engineering as well as other fields.

However, deep learning may also have a major impact in other areas of engineering and science where it has the potential to replace or augment existing carefully crafted algorithms that are designed to address various purposes. Specifically for the field of medical imaging this perspective has been laid out in [4].

In this paper we investigate the use of deep learning for the purpose of sinogram completion in CT, which has immediate applications to metal artifact reduction, but may also be used to address the effects of projection data truncation, etc.

MAR (metal artifact reduction) in CT has been a field of study for more than three decades (see [5] for an overview and additional references), but only recently some approaches have been introduced commercially. However, evaluation studies such as [6,7] show that their performance is still somewhat limited and further improvements in managing adverse effects of metal objects on image quality would be beneficial.

Sinogram completion based methods form one of the main categories of MAR approaches from the published literature, with iterative methods representing the second main group. Sinogram completion (also referred to as sinogram-interpolation, or in-painting) methods generally discard the projection data that corresponds to rays within the metal trace, and replace this "missing data" by an estimate. In an ideal case the estimated data represents a good approximation of projection data that reflects the entire shape and internal structure of the imaged object, with the exception only of the metal implant (or other metal object) itself. Specifically, structures within the object are typically represented (depending on the specific shape of the structure) by generally sinusoidal traces in the sinogram. The estimated data in the missing data region should appropriately reflect this characteristic behavior, otherwise the reconstructed image will be impacted by associated streaks or banding artifacts. In some instances, additional artifacts are created through the MAR processing that were not present in the image before correction [8].

In pure projection-based interpolation approaches the missing data is estimated based on interpolation within the sinogram domain (see, e.g., [9,10,11]), while some other sinogram completion approaches utilize an initial reconstruction (maybe using a few iterations) to produce a first estimate of the structure of the imaged object which (after re-projection) helps in obtaining an improved sinogram interpolation [12,13,14].

In the approach that is presented in this paper we attempt to estimate missing data in the sinogram itself without employing an initial reconstruction step. Similar to the approach taken in other pure sinogram-based interpolation schemes we estimate in the current implementation the missing data for a single view (or a small set of adjacent views) from a detector region that is adjacent to the missing data region (i.e., from data corresponding to detector channels that are adjacent to the

B. E. H. Claus, Y. Jin and B. De Man are with GE Global Research Center, Niskayuna, NY 12309, USA (e-mail: claus@ge.com). L.A.Gjesteby and G. Wang are with RPI (Rensselaer Polytechnic Institute), Troy, NY 12180, USA.

DOI: 10.12059/Fully3D.2017-11-3110004

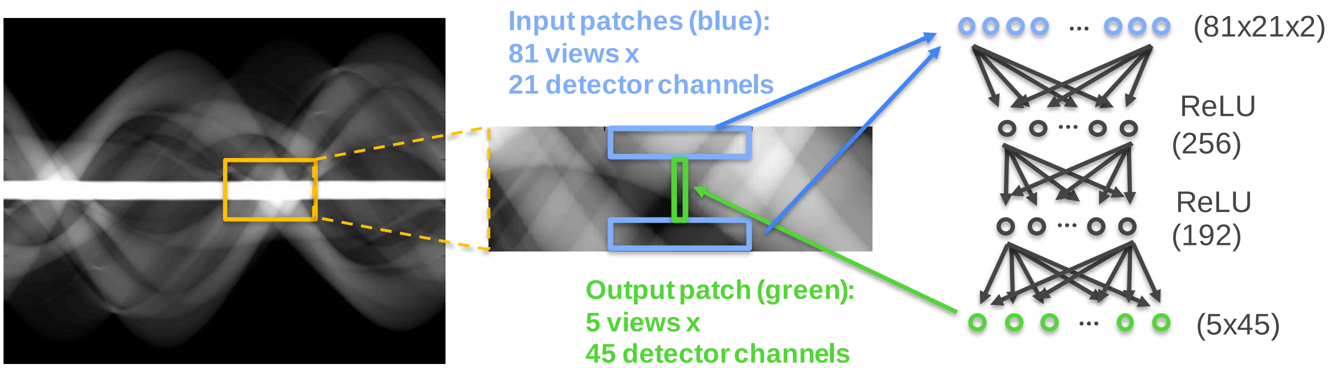

Figure 1: Example sinogram with the metal trace shown as a white horizontal band at the center (left). At the center an ROI of a sinogram is shown, with data regions corresponding to the input (blue) and the output (green) of the deep learning network marked by rectangles. The diagram on the right depicts the architecture of the used network with two hidden layers.

missing data region on both sides), and from views corresponding to an angular interval around the current view angle. This estimation process can be implemented in a straightforward way as a simple fully connected neural network.

This type of approach is reminiscent of recent work in single image super-resolution where the goal is to estimate a high-resolution image from a down-sampled/low-resolution version of the same image. In [15] results that are superior to more traditional techniques have been obtained by a simple convolutional neural network (CNN) consisting of a set of analysis filters (as the first layer), followed by a mapping of the resulting feature maps into a mapped feature space (as a second layer), which is then followed by a second convolution with appropriate "synthesis filters" and summation of the resultant images (as a third and final layer).

Inspired by this work the network in our initial approach as presented here also contains very few layers, where the first layer can be interpreted as an extraction of image features (which in our application are extracted from regions of the sinogram that are located adjacent to the missing-data region to be estimated), followed by a mapping of features and a "synthesis" of the missing data from the mapped features as the last layer.

II. Material and Methods

The sinograms used as training and validation data used in our work were generated using the Radon transform (i.e., simple line-integrals in a parallel-beam configuration) of simulated 2D objects consisting of superimposed ellipses with random orientation, size, aspect-ratio, and attenuation. For simplicity, the trace of the metal is assumed to be a band of constant width at the center of the detector, as would be created by a circular metal object at the center of the image region. (In the sinograms shown in this paper, the metal trace corresponds to a horizontal band in the sinogram.) Consequently, a central circular region in the original simulated image is replaced with the local mean value (with a smooth transition at the edges of the circular region) prior to creating the simulated projection data that is used for training. In this manner we avoid the projection data in the missing data region to contain image information that is due to structures that are located in the metal region of the object (and which therefore should not be used for training).

Images are created with a size of 511x511 voxels, and a sinogram is created for view angles spanning 360 degrees, with a 0.5 degree separation between projections, resulting in a sinogram consisting of 720 views. The images contain a simulated circular metal region of diameter 45 voxels at the center of the image, and the missing data region in the resultant sinogram corresponds to a (horizontal) band of 45 detector channels. Image simulation as well as creation of projection data and reconstructed images was performed using Matlab's radon and iradon functions.

The deep learning network takes as input two patches of size 81x21 in the sinogram, with one patch located on either side (top and bottom) of the missing data region. Each patch corresponds to an interval of 21 detector channels adjacent to the missing data interval on the detector, covering an angular range of +/-20 degrees relative to the considered view angle (i.e., from 40 views before to 40 views after the currently considered view angle). The corresponding output patch corresponds to an interval of 5 views (from two before to two after the current view angle), spanning the entire height of the metal trace (i.e., 45 detector channels). This geometry is illustrated in Figure 1, where a sinogram is shown (left), with a metal trace indicated by the horizontal white bar across the sinogram. An ROI of the sinogram (enlarged) is also shown, with the input patches (for a given considered view angle) indicated by blue rectangles, and the corresponding output patch is marked as a green rectangle. Figure 1 (right) also illustrates the network architecture used in our work, where in the first layer a set of 256 features (each corresponding to an analysis filter of size 81x21x2) is extracted, which is then mapped to a set of 192 features, where each of those mapped features is then propagated to the target patch in the missing data domain by using a 5x45 "synthesis" patch. With the

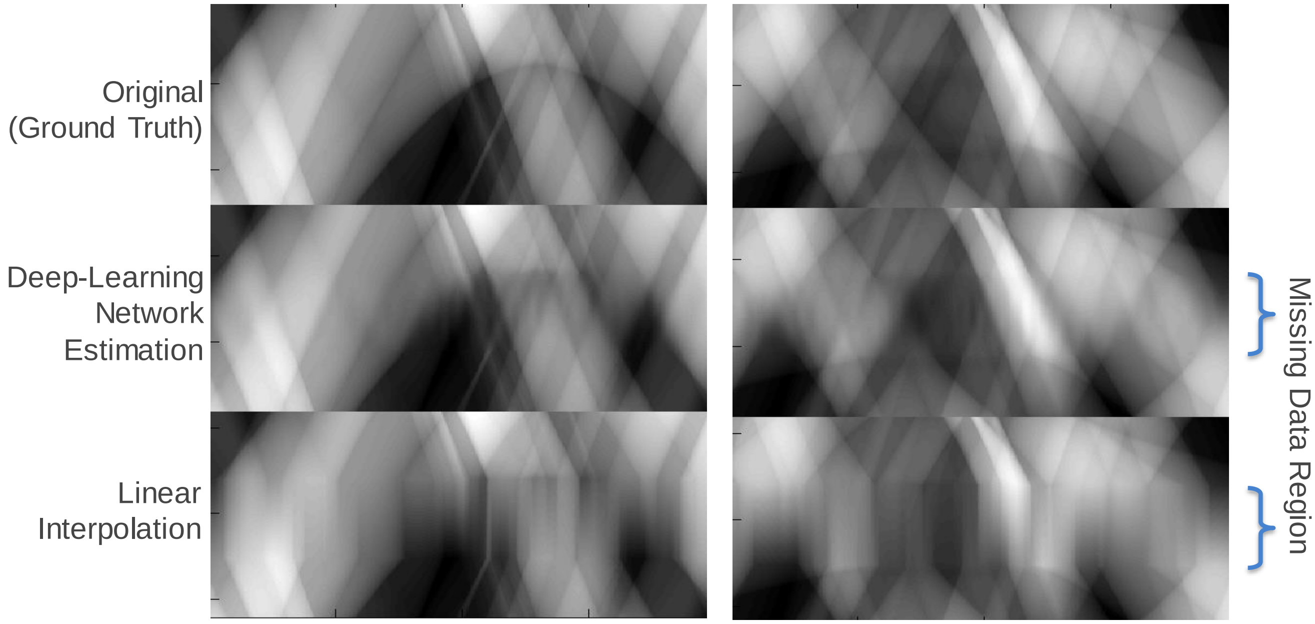

Figure 2: Example ROIs of two sinograms. Each of the two images shows three subimages: a central section of the original (ground truth) sinogram at the top, the corresponding estimated (by deep learning) sinogram at the center and for comparison the sinogram with linear interpolation of the metal trace region at the bottom. Each subimage has a height of 87 pixels and contains the metal trace (horizontal band with a height of 45 pixels) at the center, and the adjacent bands (of 21 pixel height each) on which the estimation is based immediately above and below. Please note how the interpolation using the deep learning network performs much better in capturing and representing the sinusoidal character of the traces due to the structures of the imaged object.

exception of the last layer, each layer uses a ReLU (rectifying linear unit) non-linearity.

The training of the network was based on about 30,000 datasets that were extracted from a set of 500 simulated objects and the associated sinograms.

For evaluation, interpolated sinograms were generated from individual patch-based estimates by creating an estimate for each single view angle (using the trained deep learning network), and then performing a simple averaging of the resulting overlapping output patches.

III. Results

In Figure 2 we show results of ROIs of interpolated sinograms. Specifically, the central region of a sinogram containing the missing data region (a horizontal band with a height of 45 detector channels) as well as the input data regions (bands above and below the missing data region, respectively, each with a height of 21 detector channels) is shown. For better visual evaluation, each interpolated dataset is shown in the context of the corresponding "ground truth" dataset (top), as well as a sinogram generated with simple view-by-view linear interpolation across the missing data interval (bottom). We can see that the sinogram interpolation created with the deep learning network achieves a much better "blending" with the adjacent known data bands, and is clearly superior in capturing and representing the characteristics of the data consisting of "superimposed sinusoidal traces".

In addition, for sinograms interpolated with our deep learning approach the reconstructed image was reconstructed. For reference, we also show the original image as well as a reconstruction obtained from a sinogram interpolated with simple view-by-view linear interpolation. These results are shown in Figure 3, where the top row shows the original image (left), the image reconstructed from the deep learning network-interpolated sinogram (center), as well as the image based on a linearly interpolated sinogram (right). The corresponding difference images (original – deep learning) and (original – linear interpolation) are shown in the bottom row (center, and right, respectively). Note the obvious reduction in streaking and banding artifacts for the image reconstructed from the deep-learning network-interpolated sinogram. This is also reflected in a reduction of the RMS (root mean-squared error) of 37% (RMSE error of 0.039 for reconstruction from deep-learning network interpolated data vs. RMSE error of 0.063 for reconstruction from view-by-view linear interpolation across missing data region). The remaining artifacts seem to be mostly of a high-frequency nature, which may be due to limitations in the number of levels and number of neurons in the deep learning network.

IV. Discussion

The sinogram completion results that are presented in this paper demonstrate the potential of deep learning techniques to achieve good performance in addressing streak and banding artifacts due to metal in the field of view, with potential applicability to other missing data scenarios (e.g., truncation).

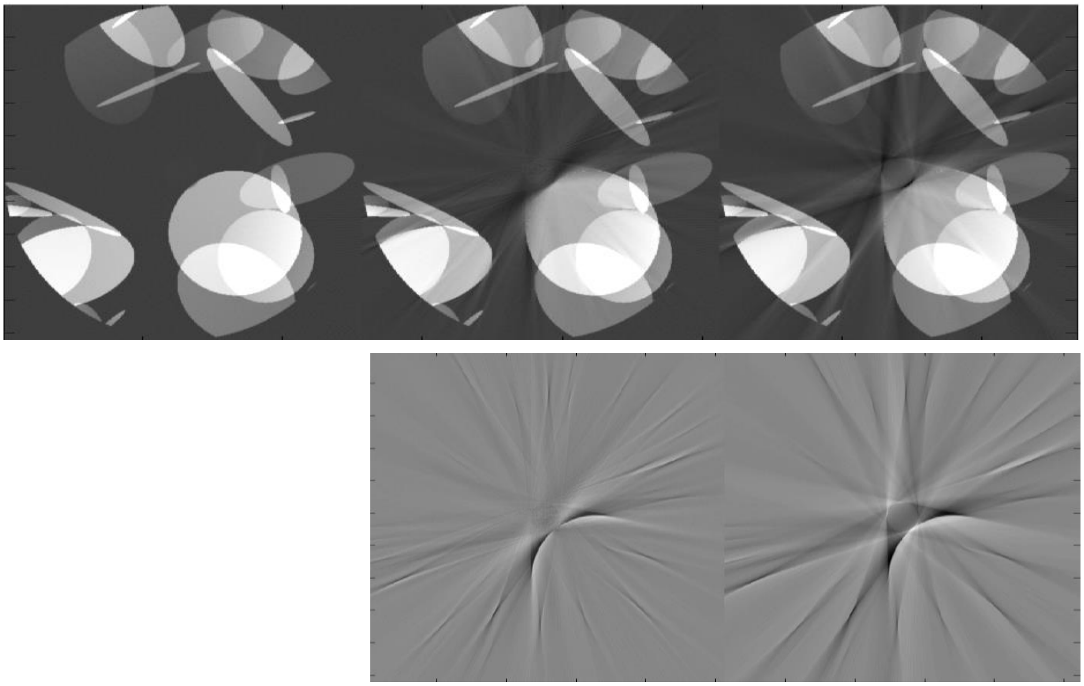

Figure 3: Example images: original (top-left), reconstruction after deep learning interpolation of metal trace (top-center), and reconstruction after linear interpolation of metal trace region (top-right). The bottom row shows the difference of the respective reconstructions above relative to the original data. Please note the clear improvement of the banding and streaking artifacts as displayed for the image reconstructed from the deep learning interpolated sinogram, corresponding to a RMS error improvement of 37%. The remaining artifacts in the result obtained with our approach seem to be predominantly of a high-frequency nature, which may be due to the relatively small number of layers and neurons in our network.

Experiences from the deep learning community suggest that increasing the number of layers in the neural network has the potential to drastically improve performance (even when the total number of neurons is decreased). Indeed, for the single-image super-resolution problem [15] that served as inspiration for the initial design of the network that we used here, a "very deep" convolutional network with dramatically improved performance was developed [16]. The performance achieved in our experiments may be limited further by the relatively small size of the input data patches. Further investigation and development of appropriate network architectures will address these limitations.

Generalization of our approach toward potential applicability in practical patient scanning scenarios will also need to address the numerous simplifying assumptions that we made in this work. For example, the fact that the metal trace is assumed to be a straight horizontal band through the sinogram may be addressed by appropriately re-binning the projection data. Also, as currently implemented, the sinogram interpolation assumes a metal trace of a fixed and constant width. This may be addressed by appropriate interpolation, down- and/or up-sampling, and other processing of the sinogram. Obviously, the case of multiple metal objects, where the traces intersect in the sinogram domain will also have to be addressed.

Furthermore, as the trained deep learning network will reflect not only the pure missing data related limitations and artifacts but also the characteristics of the training data itself, the use of appropriate anatomically correct simulated data, and, ultimately, patient data (or simulations generated based on patient data) will be used for training and evaluation of the deep learning network.

REFERENCES

[1] Krizhevsky A, Sutskever I, Hinton G. 2012 ImageNet classification with deep convolutional neural networks. In Proc. 26th Annual Conf. on Neural Information Processing Systems, Lake Tahoe, NV, 3–6 December 2012, pp. 1090–1098.

[2] Hinton G et al. 2012 Deep neural networks for acoustic modeling in speech recognition. IEEE Signal Process. Mag. 29, 82–97.

[3] Yann LeCun, Yoshua Bengio & Geoffrey Hinton, Deep learning, Nature 521, 436–444 (28 May 2015).

[4] Wang, G. "A Perspective on Deep Imaging", IEEE Access, vol 4., November 2016.

[5] Gjesteby, L., De Man, B., Jin, Y., Paganetti, H., Verburg, J., Giantsoudi, D., Wang, G., Metal Artifacrt Reduction in CT: Where Are We After Four Decades?, IEEE Access, vol. 4, 2016.

[6] J. Y. Huang et al., ''An evaluation of three commercially available metal artifact reduction methods for CT imaging,'' Phys. Med. Biol., vol. 60, no. 3, p. 1047, 2015.

[7] Wagenaar D, van der Graaf ER, van der Schaaf A, Greuter MJW (2015) Quantitative Comparison of Commercial and Non-Commercial Metal Artifact Reduction Techniques in Computed Tomography. PLoS ONE 10(6), 2015.

[8] J. Müller and T. M. Buzug, ``Spurious structures created by interpolationbased CT metal artifact reduction," in Proc. SPIE 7258, Med. Imag., Phys. Med. Imag., Mar. 2009.

[9] Kalender W. et al. "Reduction of CT artifacts caused by metallic impants." Radiology, 1987.

[10] Glover G. and Pelc N. "An algorithm for the reduction of metal clip artifacts in CT reconstructions." Med. Phys., 1981.

[11] Mahnken A. et al, "A new algoritbm for metal artifact reduction in computed tomogrpaby, In vitro and in vivo evaluation after total hip replacement." Investigative Radiology, vol. 38, no. 12, pp. 769-775, Dec. 2003.

[12] Meyer E, Raupach R, Lell M, Schmidt B, Kachelriess M., Normalized metal artifact reduction (NMAR) in computed tomography. Med Phys. 2010 Oct;37(10):5482-93.

[13] Boas FE, Fleischmann D., Evaluation of two iterative techniques for reducing metal artifacts in computed tomography. Radiology. 2011 Jun;259(3):894-902.

[14] Meyer E, Raupach R, Lell M, Schmidt B, Kachelrieß M. Frequency split metal artifact reduction (FSMAR) in computed tomography. Med Phys. 2012;39: 1904–16.

[15] Dong C., Loy C.C., He, K., Tang, X., Image Super-Resolution Using Deep Convolutional Networks, IEEE Trans. PAMI, 38(2), 295-307, 2016.

[16] Kim, J., Lee, J.K., and Lee, K.M., Accurate Image Super-Resolution Using Very Deep Convolutional Networks, IEEE CVPR2016, 1646-1654.