Multidomain CT Metal Artifacts Reduction Using Partial Convolution Based Inpainting

Artem Pimkin*, Alexander Samoylenko†, Natalia Antipina‡, Anna Ovechkina§,
Andrey Golanov‡, Alexandra Dalechina¶ and Mikhail Belyaev*
*Skolkovo Institute of Science and Technology, Moscow, Russia
Email: {a.pimkin, m.belyaev}@skoltech.ru
†Kharkevich Institute for Information Transmission Problems,
Moscow Institute of Physics and Technology, Moscow, Russia
Email: samojlenko.ai@phystech.edu
‡N. N. Burdenko National Medical Research Center of Neurosurgery, Moscow, Russia
Email: {nantipina, golanov}@nsi.ru
§Lomonosov Moscow State University, Moscow, Russia
Email: av.ovechkina@physics.msu.ru
¶*Moscow Gamma Knife Center*, Moscow, Russia
Email: adalechina@nsi.ru

Abstract—Recent Metal Artifacts Reduction (MAR) methods for Computed Tomography are often based on image-to-image convolutional neural networks for adjustment of corrupted sinograms or images themselves. In this paper, we are exploring the capabilities of a multidomain method, which consists of both sinogram correction (projection domain step) and restored image correction (image-domain step). We formulate the first step problem directly as sinogram inpainting, which allows us to use methods of this specific field, such as partial convolutions. Moreover, we propose a synthetic data generation pipeline to avoid problems with overfitting to metal shapes set and an artifacts formation technique. The proposed method achieves state-of-the-art (-75% MSE) improvement in comparison with a classic benchmark - Li-MAR.

Index Terms—Convolutional Networks, Computed Tomography (CT) images, Metal Artifacts Reduction, Sinogram Inpainting, Partial Convolutions

I. INTRODUCTION

A. General

Computed Tomography (CT) is a commonly used imaging method in disease diagnosis and treatment planning. In particular, dose distributions in radiation therapy are calculated as the solution of a forward problem; CT images determine the electron density of the irradiated tissues and patient-specific anatomy. High-density objects (e.g., containing metal) may occur in the area of interest and strongly affect the attenuation of X-Rays that may lead to distortion of the final image reconstructed from an inconsistent sinogram [1]. These image artifacts could have a significant impact on the dose calculation accuracy and reduce the visibility of organs and structures close to the metal objects [2], [3].

B. CT artifacts in Brain Radiosurgery

Though CT artifacts may affect the quality of radiation therapy in various clinical scenarios, we focus on cerebral arteriovenous malformations (AVMs), which are the focal conglomerations of the pathological vessels in the brain. The primary goal of the AVM management is to prevent the risk of intracranial hemorrhage [4]. AVMs treatment options include microsurgery, radiosurgery, embolization or combination of these modalities [5]. Within our study, we analyzed CT scans of the patients with AVMs who underwent prior embolization to block blood circulation within pathological vessels before radiosurgery.

Embolic agents injected during the first stage of this treatment can be divided into solids (for example, metal coils) and liquids. Metal components of the embolic agents, as well as coils' metal body, cause severe artifacts on CT scans, producing both bright regions of high absorption and dark regions of low absorption. Several studies have proposed that the presence of these high-density objects induces beam hardening artifacts on CT scans and might distort the dose calculation accuracy of the radiosurgery planning [6]. Figure 1 shows CT scans of two patient with AVMs after embolization (see more technical details on types of embolization in Section III-A).

The artifacts need to be identified during treatment planning and somehow fixed to avoid possible dose calculations errors [7]. In current clinical practice, this procedure is usually manual: doctors segment primary tissues like brain or skull and then replace electron densities of corrupted areas by the electron densities appropriate to the tissue. This time-consuming and a rather rough operation must be repeated for each affected slice of the CT image (usually 50-100 2D images for one patient).
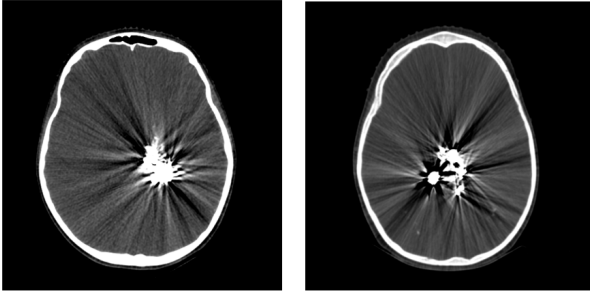
Fig. 1. Liquid (left) and solid (right) embolization of the AVMs with notable artifacts on the CT images due to the metal components of the embolic materials.

C. Problem

Thus we formulate a Metal Artifacts Reduction problem as the following: given an input 3D tensor that represents a CT scan corrupted by the presence of high-density objects, generate a corresponding CT image with suppressed artifacts. Due to the high importance of the MAR, a variety of different methods has been developed during the last 40 years; see an overview [8]. It's important to note that in addition to usual image space, each CT slice can be naturally represented as a sinogram, which mathematically is the direct Radon transform of the same slice in the image space. Indeed, the majority of the proposed solutions can be divided into two large groups based on the domain of the input data:

1) Algorithms consisting of removal of the high-density area from sinogram with further reconstruction based on the uncorrupted parts, also known as projection-based, e.g., classical Li-MAR which consists of linear interpolation for missing data within the metal trace. [9]. Nowadays, this problem may be solved using image-to-image deep convolutional networks (e.g., [10]).
2) Image-based solutions that use image-to-image networks and reduce artifacts directly on the pixel data(e.g., [11]).

D. Contribution

First, we propose a deep learning-based method that combines both approaches described above: it consists of two models that process the image representation in two domains. The first inpainting model is responsible for the removal of the distorted metal trace from the sinogram. The second refining model corrects the residual artifacts after image restoration. In this work, we have successfully verified the following statements:

1) Refining model improves the quality of the result since even minor inconsistencies in the sinogram may lead to significant artifacts on the restored image.
2) Sinogram adjustment via the inpainting model may simplify the problem for the direct image-to-image refining model.

And while the idea of the dual-domain method already appeared in some works (e.g. [12]), we formulate the problem of restoration of the corrupted area of sinogram directly as image inpainting. It allows us to use state-of-the-art approaches for this step, i.e., partial convolution-based neural network proven to outperform classic fully-convolutional end-to-end approach (method was introduced in [13] for inpainting of irregular holes).

Second, we propose a new pipeline for random generation of high-density objects that allows us to train the networks in a self-supervised setting due to the absence of (artifact-free CT, CT with artifacts) pairs. Our method doesn't rely on simple artifacts generation in the image domain. This is an important part of the proposed method since it allows us to avoid two major problems of the common approaches of the synthetic datasets for artifacts reduction: (1) mismatch between simple generation algorithms and the diversity of the shapes on real CT scans of AVMs patients and (2) possible overfitting to the structure of the simulated artifacts (see more details in Section II-B).

We compared the performance of the proposed model with FCN-MAR [14], Deep-MAR [15] and CNN-MAR [16] as one of the most popular and recent deep learning based methods of metal artifact reduction. The results show the superior quality of our method.
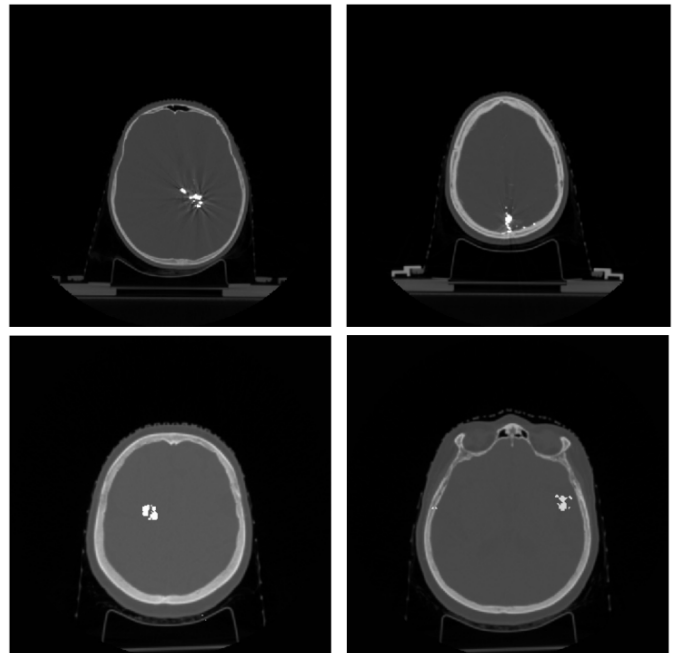

Fig. 2. Example of real shapes of high density objects (above) and shapes generated via proposed pipeline (below).

II. PROPOSED SOLUTION

The main idea of the proposed solution is to combine both image-based and projection-based approaches. The first step is the removal of the metal trace from the sinogram with the restoration of deleted areas (projection domain step). The second step is the elimination of residual artifacts from
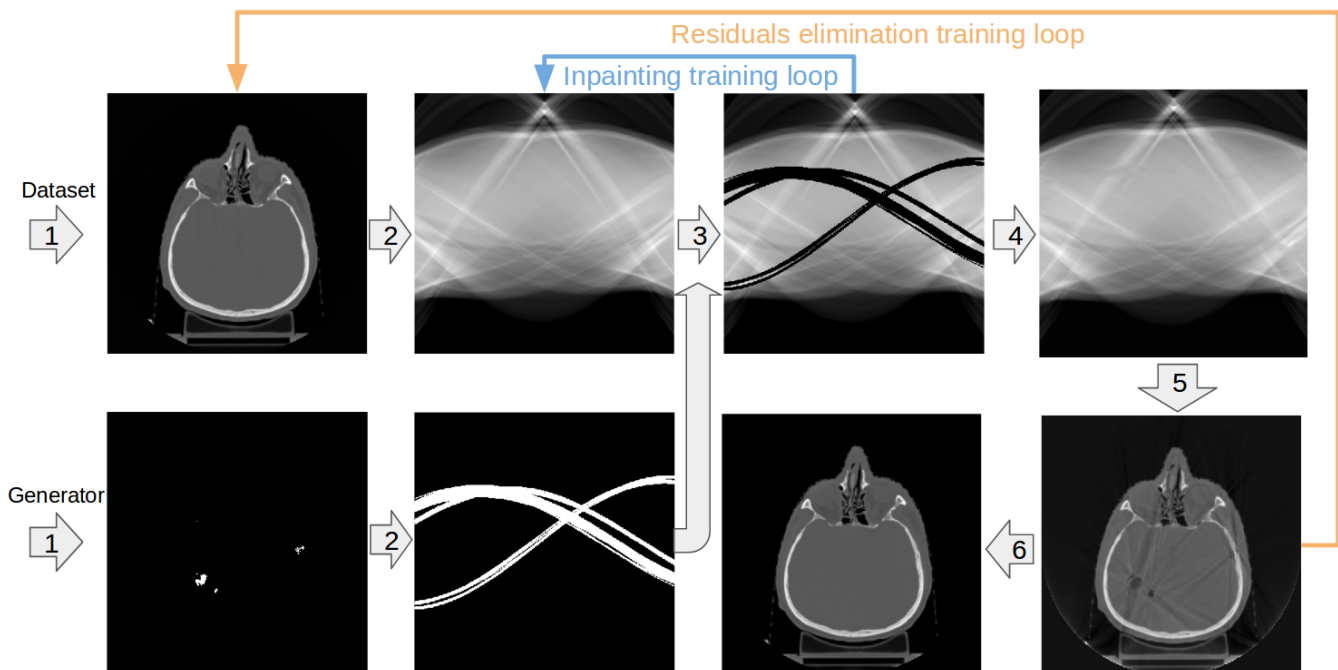
Fig. 3. Overall training pipeline. Step 1: random generation of high-density objects. Step 2: sinogram calculation for (a) input CT image an (b) generated objects. Step 3: removing parts of sinogram (a) using sinogram (b) as a binary mask. Step 4: the output of the inpainting model (the blue arrow highlights training input-output pairs). Step 5: inverse Radon transform of (4). Step 6: suppression of the residual artifacts by the image-to-image network (the orange arrow represents training data).

the restored image (image domain step). A more accurate formulation of the pipeline is further in Section II-A.

A. Overall structure

Firstly, it is important to mention that we implemented slice by slice pipeline due to avoidance of spatial data inconsistencies, e.g., different spacing between slices, which is common for medical imaging. Thus, we propose the following algorithm structure for each of the corrupted slices:

1) Cut a mask of the high-density object using threshold (since CT voxel intensities represent its density). It's a common method to detect such objects (e.g [15]).
2) Use Radon transform to obtain sinograms (a) for the corrupted image itself and (b) for the mask of the found high-density objects.
3) Remove these parts of sinogram (a), which corresponds to the sinogram (b), i.e., affected by the high-density objects.
4) Restore cropped sinogram area using an inpainting network (see details below). Ideally, it has to match the sinogram of the image with the absence of both high-density objects and artifacts caused by it.
5) Calculate the image from the modified sinogram using the inverse Radon transform.
6) Adjust the image to remove residual artifacts using the second, refining networks (see details below).
7) Add the high-density object to the final image using the mask from step 1.

At steps 4 and 6 we use convolutional neural networks as described in Section II-C. To measure the effectiveness of the proposed algorithm in comparison with only sinogram inpainting and direct image to image model, we also trained the same pipeline without step 4 and step 6, respectively. More details on the pipeline and training loops can be found in Figure 3.

B. Synthetic data

Due to the absence of the paired images, we decided to create an algorithm to generate realistic high-density objects and train networks in self-supervised settings. We used CT scans of the patients with no such high-density objects and related artifacts to create a synthetic dataset. We built a random 3D-shape generator such that generated objects had a similar structure as real high-density objects on head CT scans. Indeed, as we mentioned in the introduction, we aimed to solve two common issues with algorithms for the generation of the synthetic datasets for artifacts reduction:

1) Shapes diversity on a real scan data. Common method of high-density shapes generation is creating a dataset of objects from the real scans (e.g. [12], [14]), which is not so applicable to the field of radiosurgery due to wide range of the shapes of AVMs after endovascular embolization. That is why for this specific case of metal objects appearing in radiosurgery practice, we suggest using a generator of a structurally close to the real, but still random and diverse shapes, which leads to the increased robustness of the model.

2) Artifacts characteristics overfitting. MAR algorithms that operate in the image domain are indirectly relying on the structure of the artifacts (e.g., [16]). However, as those artifacts are simulated, it may introduce a systematic bias and a subsequent drop in models' performance on real artifacts. Thus, we suggest using only the binary mask of high-density objects with no utilization of an image with primary artifacts.

To achieve this, we propose the following algorithm:

1) Select a volumetric range to place objects randomly (uniformly from 0 % up to 10% of the linear size of the image).
2) Put a random number (uniformly from 1 to 25) of geometrical structures (ball, octahedron, parallelepiped) of linear size up to 10 pixels into this volume.
3) Merge these random structures using the morphological closing operation to obtain a randomly shaped object or a small set of them.
4) Put into the volume up to 30 geometrical structures of small size (from 1 to 3 pixels) to obtain an object with outliers.
5) Select a position of the obtained object randomly (via uniform 2D distribution) so that the overlap between the mask and the brain is $\geq 95\%$.
6) Repeat the process up to 10 times to obtain a scan with multiple objects

Figure 2 shows examples of real and generated objects. Using that algorithm for each CT image, we created 30 differently distributed object masks with an approximate depth of 90 slices per sample on average.

C. Network architectures

For deep learning problems formulated above, we used the following architectures:

1) **UNet [17]** - a well known fully-convolutional network with additional connections between encoder and decoder. As it was originally proposed for small medical datasets, Unet is also performed well in the task of metal artifacts reduction on sinograms (e.g. [10]) as well as in the direct image to image artifacts removal process (e.g. [11]). We use this architecture for step 6 of our algorithm (reduction of the residual artifacts on the images).
2) **UNet with partial convolutions** - this architecture differs from the UNet mentioned above with convolutions that were replaced by partial convolutions [13] which are masked and renormalized to be conditioned only on the pixels that are not masked. In this work, we formulate the problem of sinogram correction directly as inpainting, which allows us to integrate this state-of-the-art method to our pipeline directly.

III. EXPERIMENTAL SETUP

A. Data

Our data consisted of two datasets from Radiosurgery Department at the Burdenko Neurosurgical institute. To obtain CT scans without artifacts, we selected 163 CT scans of the patients treated with radiosurgery for trigeminal neuralgia (TN). All these patients underwent computed tomography as a part of the planning process. For the acquisition of the CT images, GE Optima 580 was used. The scanning parameters were as follows: 120 kVp, 350 mA, FOV 300 mm, slice thicknesses 1,5 mm.

Each scan is represented by a 3D tensor of shape $512 \times 512 \times N$, where N is the number of axial slices varying from 130 to 210. The data was divided into three parts patient wisely: a training set for the first model, a training set for the second model (approx. 10000 512×512 images for both training sets), and an independent testing set (approx 9000 512×512 images).

Also, we selected 47 CT scans of the patients with AVMs that were treated with both radiosurgery and prior endovascular embolization or surgery between 2014 and 2019. Embolic agents can be divided into solids (for example, metal coils) and liquids. Two commonly used liquid materials for cerebral AVM embolization are Onyx (ethylene-vinyl alcohol copolymer (EVOH))and n-BCA (n butyl cyanoacrylate) [7]. The n - BCA agent is a fast - polymerizing liquid adhesive. It includes iodine (Z = 53) as a visualization material during the injection. Onyx contains suspended micronized tantalum (Z = 73) powder. As we discussed, metal components of the embolic agents cause severe artifacts on CT scans, producing both bright regions of high absorption and dark regions of low absorption. This dataset was used in two ways

1) for the qualitative validation of the proposed solution by experts;
2) study the shapes and sizes of real metal-containing objects to design our method for synthetic data generation.

B. Preprocessing

All input images for both sinogram inpainting and restored CT slice residuals reduction were linearly normalized to fit into $[0, 1]$ window. Such a simple preprocessing was used to maintain the physical sense of voxel intensities.

C. Metrics

We used a standard MAE (Mean Absolute Error), MSE (Mean Squared Error), and SSIM (Structural Similarity) metrics to measure the quality of artifact reduction on the CT scans. In all cases, we compared the original image with the artificially corrupted and then restored one.

We validated our models in two ways:

1) We compared our multidomain method with image-domain only and sinogram-domain only networks, as we described in Sections II-A and II-C.
2) For comparison with other recent convolutional network based methods, we've implemented classic Li-MAR and measured relative increase of performance in terms of MSE. Due to the absence of the public MAR datasets and a large number of possible pitfalls in reimplementing architectures from other works, we examined the quality of our approach using the relative difference between model and Li-MAR performances as a metric. This is a
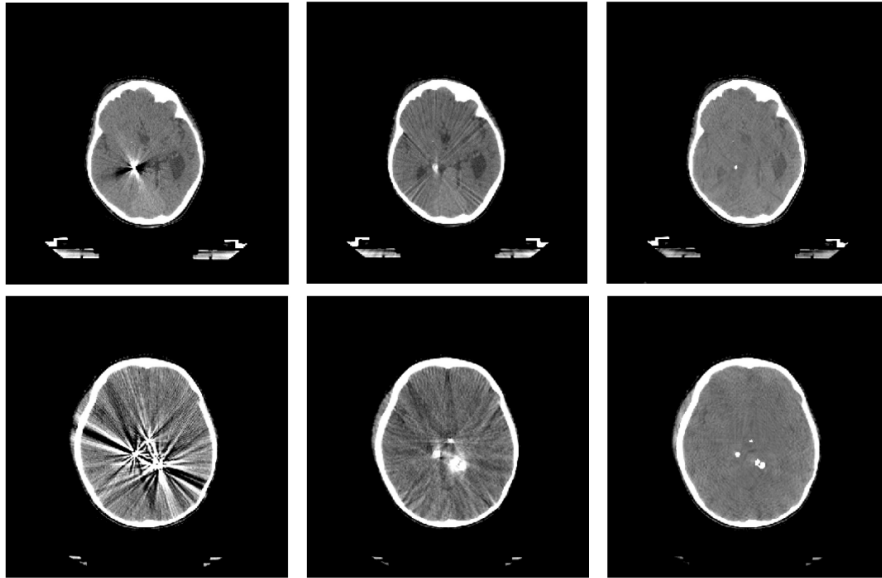
Fig. 4. Examples of processing real corrupted CT scan with intensities in brain window (40+-80 HU). From left to right: original image, Li-MAR output, the output of the proposed solution. Each row corresponds to a different patient.

common choice for the field: relative improvement of Li-MAR performance is often reported across papers (e.g., [14], [15] and [16]), which allows us to compare complex methods despite the absence of the public dataset.

D. Training

For our full pipeline and approaches comparison, we end up with training 3 models: sinogram inpainting (UNet with partial convolutions), residual artifacts elimination (UNet) and image-to-image model for the pipeline without inpainting step (Unet). All of the models mentioned above were trained using Adam optimizer with the initial learning rate of $5 \cdot 10^{-3}$.

Sinogram inpainting UNet with partial convolutions was trained for 500 epochs with the multiplication of the learning rate by 0.5 on each of the following epochs: 100, 200, 300 and multiplication by 0.1 on 400, 450, 475 and 490 epochs respectively. Models with plain UNet architecture were both trained for 200 epochs with 0.5 learning rate multiplication on 100, 150 and 175 epochs.

All models were trained on Zhores supercomputer [18] using PyTorch framework and DPipe[1] for configurations management and experiments setup.

IV. RESULTS

As we mentioned above, the classical benchmark for metal artifacts refuction methods is Li-MAR. Figure 4 shows examples of our algorithm and Li-MAR work on real CT scans of the patients with high-density objects that cause artifacts on the images. It is visible in the brain window that the model restores distinguishable brain structures well. Also, it is important to mention that the qualitative expert analysis using CT studies

[1] https://github.com/neuro-ml/deep_pipe

of patients with AVMs demonstrates that the method achieves impressive results on scans with large metal regions, which caused significant image distortion. An example is shown in Figure 5.

TABLE I
COMPARISON BETWEEN ONE STEP METHODS, LI-MAR AND TWO STEP METHOD IN TERMS OF MAE, MSE AND SSIM BETWEEN THE FINAL OUTPUT AND TEST SET. WE USED ORIGINAL PIXEL INTENSITIES IN HOUNSFIELD UNITS TO CALCULATE THESE METRICS.

Algorithm	MAE (HU)	MSE (HU^2)	SSIM
Li-MAR	26.6	3282	0.94
Inpainting-only	22.0	10170	0.94
Image-to-image only	14.5	2064	0.97
Proposed algorithm	**9.6**	**831**	**0.99**

All the test metrics of inpainting-only, image-to-image only, Li-MAR as a classic benchmark, and the proposed method are represented in Table I. Here we can see a significant (56%) decrease of MAE between the inpainting-only method and the proposed solution. Thus we can conclude that the image-to-image network can successfully remove residual artifacts and increase the quality of the joint model. We obtained quite large MSE and relatively small MAE for the inpainting model due to the inhomogeneity of its errors caused by sinogram inconsistency. On the other hand, we can see a decrease of MAE of 34% between image-to-image only method and our two-step approach. This confirms our hypothesis that sinogram inpainting as preprocessing simplifies the task for the image-to-image model. Moreover, the proposed solution significantly outperforms Li-MAR by 64% in terms of MAE.

Table II shows a relative drop of MSE reported in different papers in comparison with the reported Li-MAR score and provided us an understanding of the relatively good perfor-
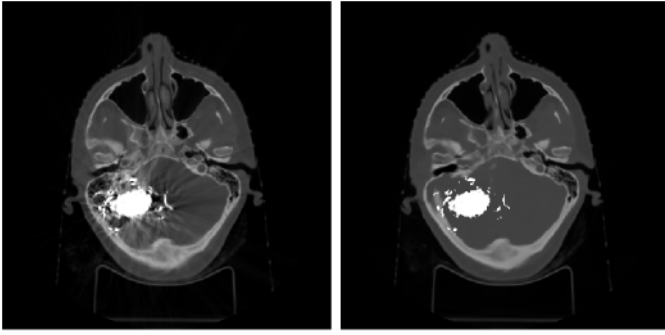
Fig. 5. An example of a CT scan with large metal-containing object causing significant distortion (left) and the output of the proposed solution (right).

mance of the proposed model. For comparison, we took three recent deep-learning-based methods operating in one or both domains:

- **FCN-MAR [14]** - method based on sinogram completion via convolutional network trained on fully synthetic data.
- **Deep-MAR [15]** - conditional generative adversarial network (cGAN [19]) based method for sinogram completion trained on fully synthetic dataset with launch on real data via transfer learning.
- **CNN-MAR [16]** - dual-domain method consisting of processing scans in image domain with different simple methods (e.g. Li-MAR) for usage as input for CNN. The output of the neural network is used to differentiate tissues, generate a prior image and, therefore, sinogram to replace metal trace in the projection domain.

TABLE II
RELATIVE DROP COMPARED TO LI-MAR PERFORMANCE.

Algorithm	MSE drop
Li-MAR [9]	0.0%
FCN-MAR [14]	-52.2%
Deep-MAR [15]	-58.5%
CNN-MAR [16]	-71.1%
Proposed algorithm	**-75.5%**

V. DISCUSSION

In this paper, we have presented a state-of-the-art multidomain deep learning procedure that consists of two main parts. Firstly, it is a metal shape generator that allows to transfer resulting model directly to the real data without using finetuning, transfer learning or domain adaptation techniques. Secondly, the method combines two CNN models: sinogram completion via state-of-the-art partial convolutions based CNN and additional adjustment after inverse Radon transform. The results of our experiments show the effectiveness of the proposed approach compared to one-step algorithms. Moreover, qualitative and quantitative analysis shows an application potential of the proposed method and its applicability, even in difficult cases. Thus, the next step of our research is going to be a clinical validation of the dose calculations for radiosurgery planning.

REFERENCES

[1] F. Boas and D. Fleischmann, "Ct artifacts: causes and reduction techniques," *Imaging in Medicine*, vol. 4, no. 2, pp. 229–240, 2012.

[2] D. Giantsoudi, B. De Man, J. Verburg, A. Trofimov, Y. Jin, G. Wang, L. Gjesteby, and H. Paganetti, "Metal artifacts in computed tomography for radiation therapy planning: dosimetric effects and impact of metal artifact reduction," *Physics in Medicine & Biology*, vol. 62, no. 8, pp. R49, 2017.

[3] D. G Kovacs, L. A Rechner, A. L Appelt, A. K Berthelsen, J. C Costa, J. Friborg, G. F Persson, J. P. Bangsgaard, L. Specht, and M. C Aznar, "Metal artefact reduction for accurate tumour delineation in radiotherapy," *Radiotherapy and Oncology*, vol. 126, no. 3, pp. 479–486, 2018.

[4] Michael T Lawton, W Caleb Rutledge, Helen Kim, Christian Stapf, Kevin J Whitehead, Dean Y Li, Timo Krings, Douglas Kondziolka, Michael K Morgan, Karam Moon, et al., "Brain arteriovenous malformations," *Nature reviews disease primers*, vol. 1, no. 1, pp. 1–20, 2015.

[5] Colin P Derdeyn, Gregory J Zipfel, Felipe C Albuquerque, Daniel L Cooke, Edward Feldmann, Jason P Sheehan, and James C Torner, "Management of brain arteriovenous malformations: a scientific statement for healthcare professionals from the american heart association/american stroke association," *Stroke*, vol. 48, no. 8, pp. e200–e224, 2017.

[6] Natan Shtraus, Dan Schifter, Benjamin W Corn, Shimon Maimon, Shlomi Alani, Vladimir Frolov, Diana Matceyevsky, and Andrew A Kanner, "Radiosurgical treatment planning of avm following embolization with onyx: possible dosage error in treatment planning can be averted," *Journal of neuro-oncology*, vol. 98, no. 2, pp. 271–276, 2010.

[7] Donald A Roberts, James M Balter, Neeraj Chaudhary, Joseph J Gemmete, and Aditya S Pandey, "Dosimetric measurements of onyx embolization material for stereotactic radiosurgery," *Medical physics*, vol. 39, no. 11, pp. 6672–6681, 2012.

[8] 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.

[9] 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.

[10] H. S. Park, Y. E. Chung, S. M. Lee, H. P. Kim, and J. K. Seo, "Sinogram-consistency learning in ct for metal artifact reduction," *arXiv preprint arXiv:1708.00607*, 2017.

[11] H. Suk Park, S. Min Lee, H. Pyung Kim, and J. Keun Seo, "Machine-learning-based nonlinear decomposition of ct images for metal artifact reduction," *arXiv preprint arXiv:1708.00244*, 2017.

[12] Wei-An Lin, Haofu Liao, Cheng Peng, Xiaohang Sun, Jingdan Zhang, Jiebo Luo, Rama Chellappa, and Shaohua Kevin Zhou, "Dudonet: Dual domain network for ct metal artifact reduction," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[13] G. Liu, F. A Reda, K. J Shih, T.-Chun Wang, A. Tao, and B. Catanzaro, "Image inpainting for irregular holes using partial convolutions," in *ECCV*, 2018.

[14] M. Usman Ghani and W. Karl, "Deep learning based sinogram correction for metal artifact reduction," *Electronic Imaging*, vol. 2018, no. 15, pp. 472–1, 2018.

[15] M. Usman Ghani and W. Clem Karl, "Fast enhanced ct metal artifact reduction using data domain deep learning," *IEEE Transactions on Computational Imaging*, 2019.

[16] Y. Zhang and H. Yu, "Convolutional neural network based metal artifact reduction in x-ray computed tomography," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1370–1381, 2018.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.

[18] Igor Zacharov, Rinat Arslanov, Maksim Gunin, Daniil Stefonishin, Andrey Bykov, Sergey Pavlov, Oleg Panarin, Anton Maliutin, Sergey Rykovanov, and Maxim Fedorov, ""Zhores"—petaflops supercomputer for data-driven modeling, machine learning and artificial intelligence installed in Skolkovo Institute of Science and Technology," *Open Engineering*, vol. 9, no. 1, pp. 512–520, 2019.

[19] Mehdi Mirza and Simon Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.