Deep Neural Networks for Ring Artifacts Segmentation and Corrections in Fragments of CT Images

Anton Kornilov*†, Ilia Safonov*†, Iryna Reimers*‡, Ivan Yakimchuk*

*Schlumberger Moscow Research, Moscow, Russia
†National Research Nuclear University MEPhI, Moscow, Russia
‡Moscow Institute of Physics and Technology (National Research University), Moscow Region, Russia
kranton94@mail.ru, isafonov@slb.com, irene.ab18@gmail.com, iyakimchuk@slb.com

Abstract—**Ring artifacts are typical defects of computed tomography (CT) that degrade the quality of a 3D reconstructed image. Existing techniques for a ring reduction have various shortcomings and limitations, in particular, a lot of them are unable to process arbitrary fragments of the image and blur artifact-free regions. We propose an algorithm for ring artifacts segmentation and reduction by deep convolutional neural networks that correct 3D fragments of the CT image by inpainting. We compare 2D and 3D architectures of networks. For the creation of a dataset with a big number of ring artifacts, we propose a procedure that is able to transfer an artifact from one image to an arbitrary place of another image. The appearance of the transferred artifact changes. For ring artifact segmentation and correction in images of sandstones and sand, the proposed networks demonstrate good visual results and outperform existing methods. The proposed technique concentrates on the Digital Rock workflow, but the networks can be adjusted for the processing of other CT images as well.**

I. INTRODUCTION

Computed tomography (CT) is a widely used method in biomedical investigations and material science. CT belongs to nondestructive techniques for obtaining information about the internal 3D structure of a substance. The spatial resolution of modern laboratory X-ray computed microtomography (microCT) systems can be about $1\,\mu m$. We process microCT images for the creation of a digital twin for various solid and granular materials for a Digital Rock (DR) physics [1]–[3], which is used in the oil and gas industry for mathematical simulation of fluids flow, as well as physical and chemical characteristics of rocks [4]. Often, we deal with a cropped fragment of a reconstructed microCT image acquired in a third-party laboratory. In this case, we have no sinograms or shadow projections of the microCT image; there is no exact information about acquisition parameters, for example, coordinates center of rotation. So, in our research, we focus on processing arbitrary fragments in the reconstructed image domain. A fragment-wise approach allows parallel processing in a cloud.

The quality of 3D reconstructed images suffers from various CT artifacts. Moreover, CT images are inherently more prone to artifacts than conventional radiographs because the image is reconstructed from a huge number of independent detector measurements [5]. Ring or circular artifacts are produced by miscalibrated or defective detector elements and look like segments of a ring centered approximately on the center of slices of a 3D image. The term *slice* denotes a 2D image that represents a cross-section parallel to the XY-plane of the 3D image in the cartesian coordinate system. Currently, methods developed by manufacturers of microCT systems are able to suppress the majority of ring artifacts, for example, by random movement between acquisitions of adjacent shadow projections [6], by filtering of shadow projections or sinograms before reconstructions, by adding the regularization to reconstruction procedure [7]. However, from several pieces to several dozen rings remain. Fig. 1a shows a fragment of slice acquired with turned off ring suppression feature in our microCT system: the entire slice area is affected by low-contrast rings. Fig. 1b demonstrates a fragment of slice acquired with the feature turned on: there is a single ring, which completely deteriorates part of the pixels. In this paper, we focus on namely such artifacts, that are rings remaining after processing by hardware and software of microCT systems. Actually, for images under consideration, such artifacts are mainly arcs; not full circles. Despite this, we call such defects rings.
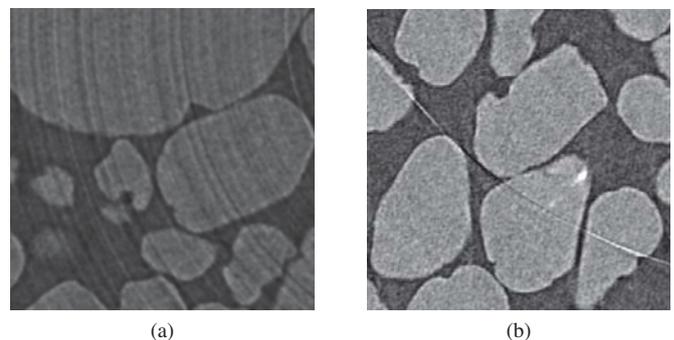
Fig. 1. Examples of fragments with ring artifacts: (a) a slice acquired with turned-off ring suppression feature; (b) a slice acquired with turned-on ring suppression feature

Usually, the Digital Rock model created from an image having a high visual quality is evaluated by the customer as a

more credible compared to model from data containing distortions. However, ring artifacts do not only degrade the visual quality of the image but also affect on the following image segmentation and mathematical simulation. Rings change the topology of pore space that, for example, acts on permeability estimation. Of course, the severity of the influence of the artifacts on the estimation of permeability depends on the nature of a specimen. Rocks with high connected porosity are almost not sensitive to rings, whereas estimation of permeability for samples with low porosity deteriorated by artifacts can have a significant bias. Let us see an example in Fig. 2. The figure shows a fragment of sandstone image affected by ring artifact and the same fragment segmented by Indicator Kriging [8], which is one of the superior segmentation techniques for microCT images [2]. One can see additional channels connecting pores. Those channels are formed by the ring artifact.
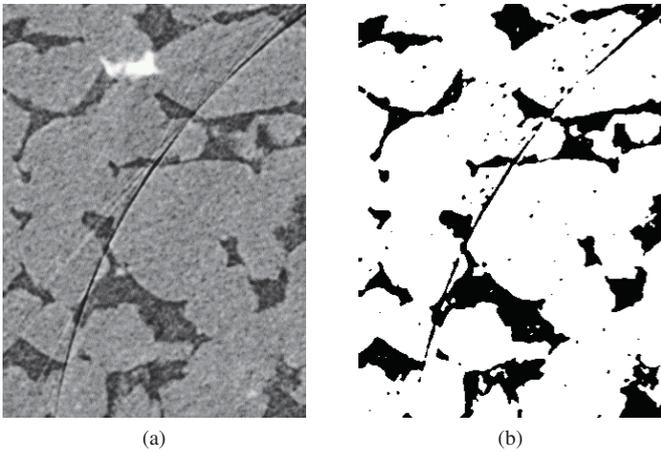
Fig. 2. Example of segmentation for DR workflow: (a) a fragment of sandstone image affected by ring artifact; (b) the segmentation result by Indicator Kriging

Existing algorithms for the rings reduction on the reconstructed images based on filtering methods affect an entire image not only deteriorated regions and require fine-tuning of parameters. It is preferable to make local corrections in automatic mode. Moreover, most ring suppression methods require exact information about the position of the center of rotation for the reconstructed image. It does not allow us to apply these algorithms in the case we have only a fragment of the reconstructed image.

The segmentation of regions deteriorated by artifacts is required for image quality estimation with aims to optimize an acquisition procedure and to select appropriate subvolume for further modeling [9].

In contrast to our previous paper [3], we propose convolutional neural networks for segmentation and local reduction of ring artifacts by nonblind inpainting. We compare 2D and 3D architectures of networks and address to problem creation of big and representative datasets containing ring artifacts. One of the main contributions of our work is a method for transferring

artifacts from one image to another. Our technique allows to generate realistic artifacts having various appearance due to the fusion of 3D ring and image in a random position of cylindrical coordinate system.

II. PREVIOUS WORKS

In this section, we review existing methods for the suppression of ring artifacts. Primarily, we focus on algorithms that are able to process reconstructed CT images. Next, convolutional neural networks (CNN) for the reduction of ring artifacts are discussed. After that, we mention several modern deep neural networks for inpainting. Finally, we describe a rings segmentation technique.

Wavelet – Fourier filtering (WFF) [10] is a popular method for suppression of ring artifacts in CT images. WFF is implemented in several software packages for the processing of CT images, for example, in TomoPy Python library [11]. This algorithm allows to suppress oriented stripes in an image. Originally WFF was applied for ring artifacts filtering in sinograms. Also, this method is used in reconstructed-image-domain for slices translated to the polar coordinates $P(\cdot)$, where ring artifacts are vertical stripes. Besides an application of WFF for pre-processing of microCT images, it is employed in DR workflow for the eliminating of a curtaining effect in slices acquired by Focused Ion Beam Scanning Electron Microscope (FIB-SEM) [12].

Firstly, the wavelet transform of the original image $I(x, y)$ is carried out to separate structural information into low frequency, horizontal, vertical, and diagonal bands $W = \{ c_{ll,m,n}, c_{hl,m,n}, c_{vl,m,n}, c_{dl,m,n} \}$ at different scales $l \in \{ 1, ..., L \}$. Then the band containing stripes (for example, vertical band $c_{vl,m,n}$ for vertical stripes) is transformed to the frequency domain with FFT. The stripes are eliminated by multiplication of the band to the Gaussian damping function in the frequency domain:

$$g(x, y) = 1 - e^{-\frac{x^2}{2\sigma^2}},$$

where (x, y) are coordinates; σ determines the standard deviation of the Gaussian filter, it is set according to the expected deviation of the stripes in the spatial domain. The final steps are inverse Fourier and inverse wavelet transformations. The main parameters of WFF filtering $F(I, L, \psi, \sigma)$ are the number of the scales L, a wavelet ψ type for the wavelet decomposition, and the standard deviation σ of the Gaussian filter.

Since WFF removes the ring artifacts from the slices in polar coordinates, it is necessary to convert them back $P_{inv}(\cdot)$ to cartesian coordinates. The conversion to polar coordinates and back can lead to a deterioration in image quality. It is preferable to use the following approach to avoid distortion. The difference between the original and filtered slice in polar coordinates is computed and translated to cartesian coordinates. The filtered slice is calculated by subtraction of this difference from the original slice according to the following expression:

$$I_f = I - P_{inv}(P(I) - F(P(I), L, \psi, \sigma)),$$

where I_f is the filtered slice; and I is the original slice containing ring artifacts.

WFF quite well suppress rings for the case turned off a rings correction feature in microCT system. As for remaining rings (as in Fig. 1b), WFF is capable to suppress such artifacts but also blurs regions of an image without ring artifacts, especially areas with circular structures suffer.

Besides WFF, there are numerous techniques for ring artifacts reduction based on filtration. For the correction of ring artifacts, the paper [13] compares two algorithms: one processes a slice in the cartesian coordinates and the other in the polar coordinates. Both methods contain the same steps: application of the median filter in the radial direction; computation of a difference image; applying a threshold for the detection of the artifacts; use of the low-pass filter in the azimuthal direction to suppress noise structures; and finally, a subtraction of the resulting image from the initial one for the artifacts reduction. The paper [14] describes an artifact reduction algorithm, which translates an image slice to the polar coordinates, detects a set of homogeneous rows within a sliding window, generates artifacts templates within a sliding window, subtracts the templates from an image in polar coordinates, and translates a slice back to the cartesian coordinates. In the paper [15], for rings reduction the following operations are declared: estimation of the local orientation in an image slice, using a representation of orientation by a structure tensor in each pixel, finding of orientations in the tensor field that correspond to circular patterns around the known rotation axis, creation of a certainty map with the probability of each pixel to contain a ring artifact; and calculation of correction image from the certainty map and the convolved original image. The paper [16] claims a method comprising of a finding of the large structures on a slice, subtracting from each structure its average intensity, a slice transformation to the polar coordinates, an application of median filter along the angular direction for small details removing, the Gaussian blurring of the slice to suppress the ring artifacts, transforming the slice back to the cartesian coordinates, and subtracting the filtered slice from the original one to reduce the ring artifacts.

All filtering-based methods blur unaffected by artifacts areas of an image. The performance of enumerated methods significantly depends on the selection of parameters. To achieve acceptable results, fine-tuning of the parameters from one image to another is required [17]. Sometimes algorithms operating in polar coordinates falsely treat circular edges as an artifact. Ideally, the edges of pores should not be located on circles centered on the axis of rotation to prevent false detection of these edges, and therefore image quality degradation. This condition is difficult to fulfill, especially for some granular materials, since the boundary between solid and void can be directly on such circles. None of the enumerated filtering methods can be applied for the processing of arbitrary fragments of a reconstructed microCT image because all of them are used information about the center of rotation.

Recently, several techniques for ring artifacts reduction based on deep learning were established. Ring artifacts cor-rection on medical CT images is based on a fusion of two 2D images: reconstructed from sinograms processed by WFF and the image from the output of a CNN [18]. On the input, the CNN takes the WFF filtered image and an image with ring artifacts. The 2D CNN has 5 sequential convolution layers and is trained with a 64×64 patches size from slices of 3D CT images. The paper [19] depicts a comprehensive model that uses two CNNs: one network that processes the image in the reconstructed image domain and the other in the projection domain. Both CNN architectures are based on 2D U-nets [20]. The method subtracts from an image with ring artifacts the output of the one from networks [21]–[23]. In both papers, authors trained models with the mean square error loss.

These methods are parameter-free at the inference stage. However, the problem of removing rings from pieces of a reconstructed image remains, since both methods use the result of sinogram processing as one of the inputs. Considered neural networks are end-to-end and do not allow to segment artifacts, whereas segmentation data is valuable. The disadvantage of these CNNs is that they work with 2D slices of 3D CT images. Since adjacent slices have valuable information for ring artifacts reduction, we suppose using a 3D CNN is preferable.

As for inpainting methods, the research in that area has been very active over the recent two decades. Image inpainting algorithms can be divided into two categories: nonblind techniques and blind ones. In the first category, the regions that need to be restored are provided to an algorithm, whereas in the second, no information about the locations of the corrupted areas is given and an algorithm automatically identifies the regions for inpainting [24]. Nowadays, deep-learning-based approaches achieve state-of-the-art results in inpainting. In our previous work [3], we proposed a blind inpainting algorithm for the ring artifacts reduction. In this work, we are interested in non-blind inpainting since we have the intention to use a mask of artifacts. There are plenty of image inpainting methods by means of CNNs [25]–[30]. The majority of analyzed publications describes generative adversarial networks (GAN) [25]–[27] and various modifications of 2D U-net [20]. Neural networks in papers [26]–[30] are intended for nonblind inpainting of relatively large rectangular regions in natural photos.

The method for segmentation of ring artifacts focusing on quality estimation of microCT images is described in [9]. The technique processes slices of 3D images one by one. The first stage is the normalization of slice intensity. Then the method computes a region of interest (ROI) of the slice by an application of a variance filter, thresholding, and morphological closing. Inside the detected ROI, the algorithm performs ring artifacts segmentation. The method translates slice to the polar coordinates and applies the matched filter to find vertical lines. Then the algorithm applies thresholding for absolute values of pixels of the filtered image and gets the result as a mask of artifacts. Next, the method uses morphological dilation with a vertically oriented structure element to merge neighboring regions in the mask. The final stages are the labeling of connected regions and elimination of regions, which have

bounding box sizes greater than the predefined parameter, and translation of the mask back to the cartesian coordinates.

There is no method, which for a fragment of reconstructed CT image allows doing segmentation and local correction of 3D regions of ring artifact remaining after conventional correction of rings in microCT system.

III. DATASETS

A. Initial images

We acquired the samples of sandstones and sand by using Bruker SkyScan 1172 microCT system with parameters described in [2]. Ring artifacts suppression feature was turned on, so, acquired images contained a relatively small number of rings, but the majority of those rings was quite noticeable. We got specimens of 8 reconstructed 3D images of the following rocks: one Unifrac sand (UFS), three Bentheimer sandstones (BHI_1, BHI_2, BHI_3); one Gravelite sandstone (GRV); one Buff Berea sandstone (BB); two Fontainebleau sandstones (FB_1, FB_2). Each reconstructed 3D image has size $3968 \times 3968 \times 1840$ voxels with a bit depth of 8 bits per voxel with a resolution of about $2.3\,\mu m$ per voxel. Fig. 3 shows fragments of these images.

To train models for the rings segmentation we need annotated patches of the images. It is necessary to have corresponding patches of the masks, where voxels related to artifacts are indicated. Initially, we segmented ring artifacts on slices of 8 available 3D images by the method from [9]. After labeling the connected components, we have 3578 3D regions in cylindrical coordinates. Some of these regions are not rings. We manually excluded erroneously segmented regions. After that, we have 2073 3D regions containing natural ring artifacts and corresponding binary masks, in which 1 denotes the voxel of a ring, and 0 denotes the voxel of a background. Regions with rings and their masks are 3D arrays in cylindrical coordinates. We made masks of artifacts convex because the segmentation algorithm left some holes and caverns. We denote sets of the regions containing ring artifacts as REG_{UFS}, REG_{BHI_1}, REG_{BHI_2}, REG_{BHI_3}, REG_{GRV}, REG_{BB}, REG_{FB_1}, REG_{FB_2}; the sets of the masks as $MASK_{UFS}$, $MASK_{BHI_1}$, $MASK_{BHI_2}$, $MASK_{BHI_3}$, $MASK_{GRV}$, $MASK_{BB}$, $MASK_{FB_1}$, $MASK_{FB_2}$. In addition, we obtain 3D binary images containing segmentation outcomes for whole initial images in cartesian coordinates and denote it as M_{UFS}, M_{BHI_1}, M_{BHI_2}, M_{BHI_3}, M_{GRV}, M_{BB}, M_{FB_1}, M_{FB_2}.

Based on M_i, we computed thickness over the Z-axis and arc length of artifacts. That information is important for the selection of the appropriate architecture of the neural network and our approach for transferring artifacts between images. Almost all rings have a thickness from 1 to 8 voxels. The distribution of arc lengths has the mean value equals 334 voxels and the standard deviation equals 282 voxels. The ring artifacts have a relatively short arc length in comparison with the maximal arc length for the entire ROI, which is equal to about $2\pi \cdot 1740$ voxels.

So, the number of regions with artifacts is small for the training of the network intended for the segmentation. Also, we are faced with class imbalance: the total number of undistorted voxels is 17705 times greater than the number of voxels of artifacts. Despite there are various solutions for the training of segmentation models on unbalanced datasets [31], [32], it is preferable to have more balanced and more diverse data.

However, for training a deep neural network for inpainting, we have a more complex problem: we need to have corresponding to each other patches of images damaged by ring artifacts and artifact-free. Where can we get such pairs of patches? According to approaches discussed in the previous section, one solution is to apply one of the methods for rings reduction before reconstruction and to reconstruct the images again. A similar approach is WFF filtering in the reconstructed-image domain and using the filtration outcome as an undistorted image. Another solution is the fusion of synthetic rings with pristine images. In the first solution, already filtered nonoriginal voxels are used as ground truth (GT). The quality of the second one depends on the adequacy of an algorithm for the formation of synthetic artifacts.

Instead of generating synthetic artifacts, we propose an algorithm for the translation of the natural rings to arbitrary places of other images. In our ring transferring procedure, the appearance of rings is altered. It allows to produce a rich and diverse dataset. Additionally, we expect that such an augmented dataset improves the performance of the segmentation since it makes the CNN more independent from the background of rings and reduces the class imbalance.

B. Translation of artifacts between images

Fig. 4 shows a scheme of artifacts translation procedure from FB_1 image to BHI_1. The top part of the figure illustrates the creation of a dataset of natural rings as described in the previous subsection. We select a region from the dataset containing natural rings by random and estimate its background: voxels of image adjacent with voxels damaged by the ring. To obtain a matrix with a "pure" artifact, we subtract the background from the region. The matrix contains signed values. Colormap in Fig. 4 shows example of range of values in the matrix after background subtraction. An artifact-free microCT image is converted to cylindrical coordinates and fused with the matrix containing artifacts with the subtracted background. The position for fusion is chosen randomly in some limitations to avoid overlapping of different artifacts. Finally, an image with the transferred ring is converted to cartesian coordinates. The width and arc length transferred ring in cartesian coordinates differ from sizes for the natural one due to the natural ring is fused with an artifact-free image in a random place of cylindrical coordinates. As a result of fusion with diverse content of microCT images translated artifacts look rather different from their source ring. Moreover, the matrix containing natural artifact with subtracted background sometimes is inverted according to the random parameter. So, our translation procedure is not just replication. One can see a
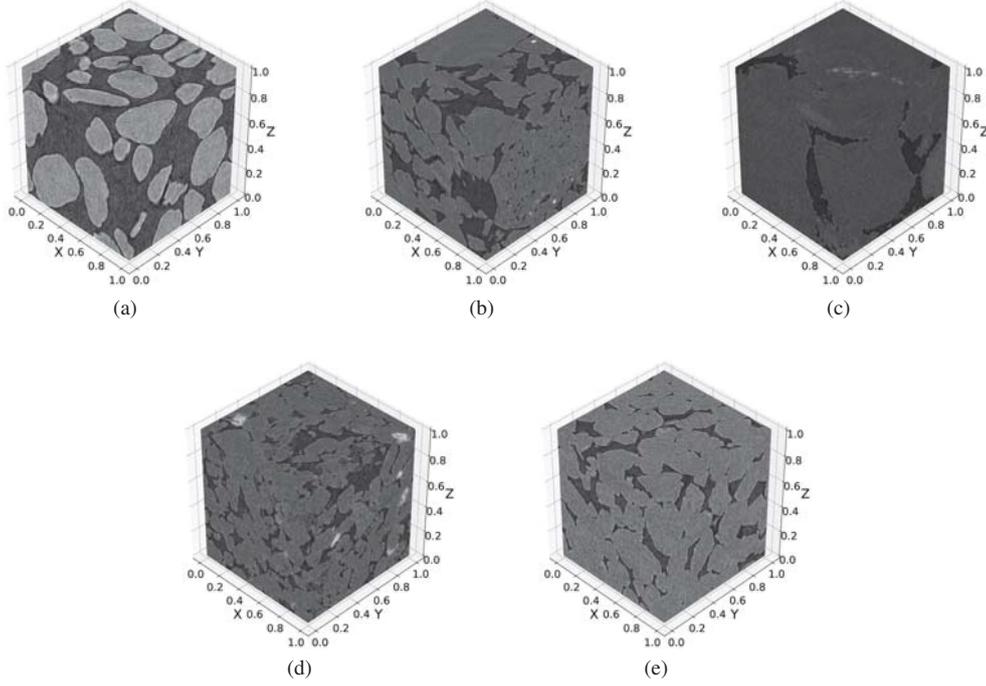
Fig. 3. Examples of 1 mm^3 cubes from images: (a) UFS; (b) BHI_1; (c) GRV; (d) BB; (e) FB_1

significant difference between the source ring and transferred artifact in Fig. 4. On the other hand, in our opinion, transferred rings look very realistic. Our rings transferring approach can be considered as special data augmentation.

The input of translation algorithm are region REG and mask $MASK$ of artifact in the cylindrical coordinates, for example $REG \in REG_{FB_1}$ and $MASK \in MASK_{FB_1}$; the image and mask in the cylindrical coordinates, on which we want to transfer the artifact, for example $P(BHI_1)$ and $P(M_{BHI_1})$. The placing artifact on the image $P(I)$ with mask $P(M_I)$ in position (x_p, y_p, z_p) is defined with the statements (1), (2), (3), (4), (5), (6), and (7). Where $x_p \in [0, M - M_a]$, $y_p \in [0, N - N_a]$, $z_p \in [0, D - D_a]$; N, M, D are the number of rows, columns, depth of $P(I)$ and N_a, M_a, D_a are the number of rows, columns, depth of REG; $x = 0, ..., M - 1$, $y = 0, ..., N - 1$, $z = 0, ..., D - 1$; y_{avg} is the number of averaged rows, in our work y_{avg} equals 20; x_{lr} is the number of columns of background from the left and right of the artifact, in our case x_{lr} equals 10, $\mu[\cdot]$ is mean value operator; $i \in \{0, 1\}$ is a random parameter for ring intensities inverting; DT is the distance transform; masks elements are equal to 0 and 1.

After transferring artifacts, the ratio of the number of background voxels to the number of voxels of rings became approximately 59. The number of regions with artifacts in total is increased to about $2 \cdot 10^5$.

C. Forming of training, validation, and testing datasets

The common practice for the creation of models by machine learning methods is to split a dataset into training, validation, and test sets. On the training dataset, we train our model. On the validation dataset, we evaluate the performance of the model during training to prevent model overfitting. On the test dataset, we evaluate the final performance.

For segmentation, GT are masks of natural and transferred artifacts. For artifacts suppression by inpainting, GT are artifact-free patches of original microCT images.

To demonstrate the stability of our approach we created four datasets: two for segmentation (see Table I) and two for reduction of ring artifacts (see Table II). Each dataset contains seven 3D images for training, one for validation and testing. UFS and UFS_r are used for testing in the *Dataset 1* and *Dataset 3*. We apply BHI_1 and $BHI_{1,r}$ for testing in the *Dataset 2* and *Dataset 4*.

To prevent overfitting, we translate rings originated from images from the training set only to images from this training set. Identically, we translate rings originated from images from the test set only to images from this set. For example, on $BHI_{1,r}$, $BHI_{2,r}$, $BHI_{3,r}$, GRV_r, BB_r, $FB_{1,r}$, and $FB_{2,r}$ we translate $\{REG \mid REG \in REG_{BHI_1} \cup REG_{BHI_2} \cup REG_{BHI_3} \cup REG_{GRV} \cup REG_{BB} \cup REG_{FB_1} \cup REG_{FB_2}\}$, and on UFS_r we translate $\{REG \mid REG \in REG_{UFS}\}$. With r index we denote images containing transferred artifacts.

IV. DEEP NEURAL NETWORKS FOR SEGMENTATION AND INPAINTING

A. CNN architectures

We need a neural network for segmentation and nonblind inpainting. It is reasonable to build a network from two parts, where the first part makes segmentation, and the second one
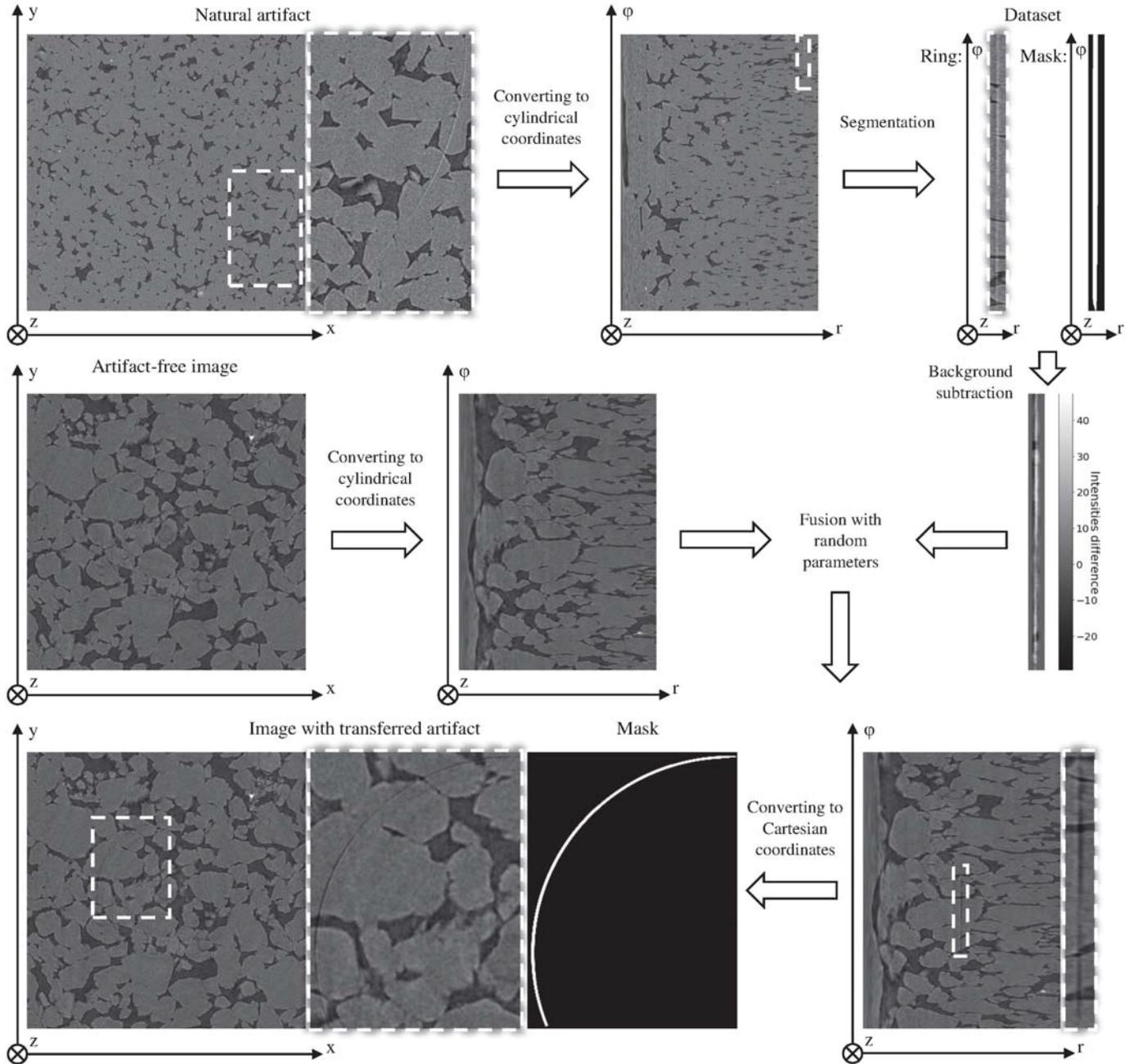
Fig. 4. Scheme of the translation procedure

employs the output of the first part for a local correction of rings. In the first part, we apply U-net [20] since this network provides pretty good results in different segmentation problems in whole and for processing of microCT images of rock samples in particular [2]. The second part is inspired by residual learning used for the reduction of metal artifacts [22] and streak artifacts [21] from medical CT images. In the concept of residual learning, a rather deep network from sequential convolutional layers predicts the so-called artifact-residual image. For making the correction, the artifact-residual image is subtracted from the image containing artifacts.

We train these two parts one by one. At first, we train part for segmentation. Then the part is frozen, and we train the second part of the network intended for correction.

The input of the first part is a patch of a reconstructed microCT image in cartesian coordinates. The output of the first part is a probability map, which elements indicated probability to be belonged in damaged by artifact area. The thresholding of the map produces a binary mask with outcomes of segmentation. For training and testing of the first part, we use patches of images with natural and translated rings as well as corresponding patches of masks containing segmentation

$$M_I^r = \begin{cases} M_I, & \begin{aligned} &x \notin [x_p, x_p + M_a - 1] \wedge \\ &y \notin [y_p, y_p + N_a - 1] \wedge \\ &z \notin [z_p, z_p + D_a - 1], \end{aligned} \\ P_{inv}(MASK(x - x_p, y - y_p, z - z_p)), & \text{otherwise}, \end{cases} \quad (1)$$

$$I^r = \begin{cases} I, & \begin{aligned} &x \notin [x_p, x_p + M_a - 1] \wedge \\ &y \notin [y_p, y_p + N_a - 1] \wedge \\ &z \notin [z_p, z_p + D_a - 1], \end{aligned} \\ I + (-1)^i \cdot P_{inv}\left(\frac{Diff(REG(x-x_p, y-y_p, z-z_p))}{Dist(x-x_p, y-y_p, z-z_p)}\right), & \text{otherwise}, \end{cases} \quad (2)$$

$$Diff(REG(x,y,z)) = \mu(REG(x,y,z)) - \left(\mu_l(x,y,z) + \frac{x}{M_a - 1} \cdot (\mu_r(x,y,z) - \mu_l(x,y,z))\right), \quad (3)$$

$$\mu(REG(x,y,z)) = \mu[REG(x,y,z), REG(x, y+1, z), ..., REG(x, y + y_{avg} - 1, z)], \quad (4)$$

$$\mu_l(x,y,z) = \mu[\mu(REG(0,y,z)), \mu(REG(1,y,z)), ..., \mu(REG(x_{lr} - 1, y, D_a - 1))], \quad (5)$$

$$\mu_r(x,y,z) = \mu[\mu(REG(M_a - x_{lr}, y, z)), \mu(REG(M_a - x_{lr} + 1, y, z)), ..., \mu(REG(M_a - 1, y, D_a - 1))], \quad (6)$$

$$Dist(x,y,z) = \begin{cases} 1, & MASK(x,y,z) = 1, \\ DT(1 - MASK)(x,y,z), & \text{otherwise}, \end{cases} \quad (7)$$

TABLE I. THE DATASETS FOR ARTIFACTS SEGMENTATION

Dataset	Training set	Validation set	Test set
1	$BHI_{1,r}$, $BHI_{2,r}$, $BHI_{3,r}$, GRV_r, BB_r, $FB_{1,r}$, $FB_{2,r}$	UFS_r	UFS, UFS_r
2	UFS_r, $BHI_{2,r}$, $BHI_{3,r}$, GRV_r, BB_r, $FB_{1,r}$, $FB_{2,r}$	$BHI_{1,r}$	BHI_1, $BHI_{1,r}$

TABLE II. THE DATASETS FOR ARTIFACTS REDUCTION

Dataset	Training set	Validation set	Test set
3	$BHI_{1,r}$, $BHI_{2,r}$, $BHI_{3,r}$, GRV_r, BB_r, $FB_{1,r}$, $FB_{2,r}$	UFS_r	UFS_r
4	UFS_r, $BHI_{2,r}$, $BHI_{3,r}$, GRV_r, BB_r, $FB_{1,r}$, $FB_{2,r}$	$BHI_{1,r}$	$BHI_{1,r}$

GT.

The input of the second part is a patch of microCT image and corresponding probability map as the segmentation result. The output of the second part is the corrected patch. For training and testing of the second part, we use patches of images with translated rings as well as corresponding patches of original artifact-free images.

Despite both parts of the network are known, for the best of our knowledge, there is no exact the same solution in the literature. We propose two deep neural networks: 2D and 3D.

Based on the distributions of the artifacts thickness over the Z-axis and the arc lengths, we set the following size of patches: 512×512 for the 2D network and $512 \times 512 \times 8$ for the 3D one.

The first part of the 2D network is the ordinary U-net with 5 levels, 64 feature maps at the top level after each convolutional layer, and doubling the number of the feature maps at each lower level. The second part has 15 sequential convolution layers that end with a hyperbolic tangent as the activation function. We choose the feature maps number that the model during training can occupy the entire GPU GeForce GTX 1080 Ti made by Nvidia.

Fig. 5 shows the scheme of our 3D network. The 3D model is almost identical to the 2D one except for the following differences. The first part of the network is based on 3D U-net [33]. For dealing with patch having different sizes along Z and other axes, we utilize an approach from [34]: the network not up/downsample feature maps over the Z-axis, on the highest resolution of the U-net consists of only 2D convolutions, and on the other scales, the first convolution is 2D and the second is 3D. In the second part, the 3D network has 13 sequential convolution layers. We had to split the network for two identical GPUs during training.

B. Training process

A common metrics for assessment of segmentation is Dice [35]. We use Dice in probabilistic form [31]:

$$Dice = \frac{\sum_{i=1}^{N} p_{1,i} g_{1,i} + \epsilon}{\sum_{i=1}^{N} (p_{1,i} + g_{1,i}) + \epsilon}, \quad (8)$$

where N is the number of the predicted voxel classes; $p_{1,i}$ is the probability that voxel i is of the ring artifact class; $g_{1,i}$ is a corresponding binary value, 0 if a voxel does not belong to
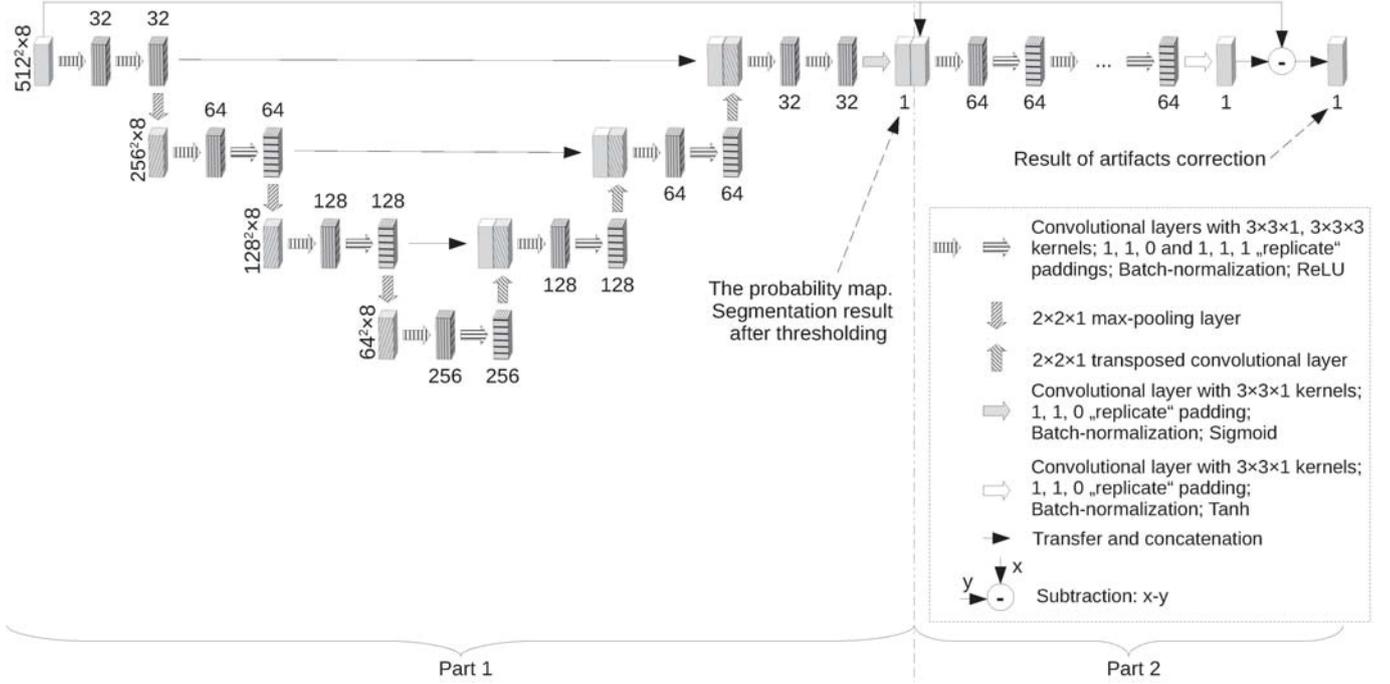
Fig. 5. Scheme of the 3D model

the class, and 1 if it belongs; ϵ is a small value for preventing division by zero.

We apply focal Tversky loss [31] for the training segmentation part of the CNN. Tversky loss is based on Dice metrics and introduces weighting coefficients α and β for weight false positives and false negatives for dealing with the segmentation of imbalanced datasets. The exponential parameter γ is used to control loss function values between background and regions of artifacts. It is used to address the issue that the loss function struggles from segmenting the relatively small regions of the artifacts.

This loss function is represented by the following formulas:

$$Tv = \frac{\sum_{i=1}^{N} p_{1,i} g_{1,i} + \epsilon}{\sum_{i=1}^{N} p_{1,i} g_{1,i} + \alpha \sum_{i=1}^{N} p_{0,i} g_{1,i} + \beta \sum_{i=1}^{N} p_{1,i} g_{0,i} + \epsilon},$$

$$L_{FTv} = (1 - Tv)^{\frac{1}{\gamma}},$$

where Tv is Tversky similarity index, $p_{0,i}$, $p_{1,i}$ are probabilities that voxel i is of the background class and the ring artifact class, respectively. The $g_{0,i}$, $g_{1,i}$ are corresponding binary values, 0 if a voxel does not belong to the class, and 1 if it belongs. In this work, we use values $\alpha = 0.7$, $\beta = 0.3$ to focus the loss to minimize false negative predictions and the suggested value $\gamma = \frac{4}{3}$ [31].

The valuable loss functions for image restoration with neural networks [36] are the weighted sum of mean absolute error (MAE) and multiscale structural similarity index measure (MS-SSIM) [37].

The loss function is:

$$
\begin{aligned}
L_{MAE,MS\text{-}SSIM}(P, GT) = \\
w \cdot MAE\left(G_{\sigma_G^M} * P, G_{\sigma_G^M} * GT\right) + \\
(1 - w)(1 - MS\text{-}SSIM(P, GT)),
\end{aligned}
$$

where P is a predicted image; GT is a ground truth image; w is the weighting coefficient; MAE is MAE, which in our work equals 0.2; $G_{\sigma_G^M}$ is Gaussian filter with the standard deviation $\sigma_G^M = 8$; and $*$ is convolution operation, $MS\text{-}SSIM$ is MS-SSIM.

Adam optimization algorithm [38] with the initial learning rate 10^{-4} is employed in training. For training the 2D network, we apply nonoverlapping patches and for training the 3D network with overlapping by 50% by Z-axis [39]. The batch sizes are equal to 4 and 1, respectively. The training is stopped when during training the evaluation metrics are almost not changed: for the segmentation case within 0.1 during the 20 iterations, for the rings reduction within 10^{-4} during the 5 iterations. After every 1000 training patches, we evaluate performance on 100 random patches from the validation dataset. As for performance evaluation metrics during training, we use the Dice for the segmentation part and MS-SSIM for the correction.

All experiments were performed on the system with two GPUs GeForce GTX 1080 Ti made by Nvidia. GPU provides 3584 stream cores, 11 GB of memory, and 11.3 Tflops of peak single-precision performance. Both neural networks were implemented via Pytorch 1.3.0. The training time 2D segmentation model is about 1 day; the 2D rings reduction

model is about 3 days; the 3D segmentation model is about 3 days; the 3D inpainting is about one week.

V. RESULTS AND DISCUSSION

A. Segmentation

For the performance evaluation of the segmentation, we use Dice metrics defined in statement (8) and the area under the precision-recall curve (PR-AUC) [40]. Table III contains the metrics for the segmentation part of our neural networks and for the method [9] based on the transformation of slices to the polar coordinates. To the best of our knowledge, there are no other techniques for the segmentation of ring artifacts remaining after conventional suppression of the artifacts in microCT system.

The metrics for the method from [9] are higher for images UFS and BHI_1 without transferred artifacts. It is explained by two reasons. Firstly, GT masks for the images were obtained namely by this method, parameters of [9] are overfitted for initial microCT images with natural artifacts. Secondly, in this case, the proposed 2D and 3D CNNs were trained only on about 2000 regions of natural artifacts. In our opinion, it is not enough training data to achieve a good generalization. For the more diverse and big datasets that included both natural and transferred artifacts, where the number of regions is about 200000, the proposed neural networks significantly outperform the existing segmentation approach. The 3D network performs better in comparison with the 2D one. The segmentation results for BHI image type are slightly better than for UFS type. We suppose that cause is presence in the training set the other image of BHI type.

Fig. 8 shows the segmentation results for a fragment from the BHI_1 image. We point to the ring artifact by a white arrow.

B. Correction

As for performance metrics for the rings reduction, we use MAE_{in} and MAE_{out} computed for voxels that belong to ring artifacts and that do not belong, respectively. Each voxel is in the range [0, 255]. MAE_{in} is the key measure, because it indicates the quality of restoration of the damaged region. MAE_{out} allows to see how big changes in unaffected by rings regions. In the ideal case, undamaged voxels should be unchanged. Also, we use MS-SSIM computed for the entire image since the paper [41] demonstrates that MS-SSIM has the highest correlation with assessments by a human. We do not expect to see a big difference in MS-SSIM for various methods, because an area of local rings is very small in comparison with another image area. Nevertheless, the visual quality of the image used for the building of the model for simulation in Digital Rock workflow is important, and we want to estimate quality both visually and numerically.

We compare our ring artifacts reduction approach with nonblind Telea's inpainting technique [42], the popular WFF method [10], and inpainting via U-net with partial convolutions [29]. Mainly, the cause to select these methods for comparison was a practical successful application of enumerated

approaches for correction of various artifacts in our laboratory before.

We apply the inpainting algorithm by Telea twice: for 2D slices of the images based on the ground truth, it is designated as *Telea 1*; and based on masks from the output of segmentation part of our 3D CNN, it is designated as *Telea 2*.

Python package *rmstripes* (https://github.com/DHI-GRAS/rmstripes) implements the WFF algorithm. For the filtering, we use the following filtering parameters: $L = 4$, ψ – Daubechies wavelet with 4 coefficients, $\sigma = 20$. We choose the parameters L and σ as minimal as possible to keep the artifact almost completely suppressed and to preserve other structures unchanged. Python package *polarTransform* (https://github.com/addisonElliott/polarTransform) allows to reduce quality loss in transformations between polar and cartesian coordinates. The angle axis in the polar coordinates is set to be two times larger than the largest dimension of the image slice in the cartesian coordinates.

The U-net model with the partial convolutions was downloaded from https://github.com/MathiasGruber/PConv-Keras. This model was pre-trained on the Imagenet dataset [43]. Since it works with three-channel images, we just copied the grayscale channel in all color channels.

Fig. 7 shows examples of the considered ring reduction methods for the fragment from Fig. 6a.

Table IV contains results for the ring reduction methods. MAE_{out} is not equal to zero for images without correction of artifacts because the proposed transferring method uses the masks of artifacts for a smooth transition of intensity using the distance transform in Equation (2); thus we get a small difference of intensities around the artifact masks. The performance of *Telea 1* and *Telea 2* are almost identical. WFF better corrects artifacts than the inpainting by Telea, but it significantly degrades the quality of other parts of images. The method based on U-net with the partial convolutions degrades the quality of the images. One can see, that in example from Fig. 7d, intensities at the ends of the artifact arc turned to black and some structures near the artifact have deteriorated shape. However, it is worth remembering that this model was trained on the Imagenet dataset. We assume, if it had been trained on our dataset, the result may improve. The proposed CNNs demonstrate the best performance. The 3D CNN has the highest values for all metrics except the MAE_{out} for *Dataset 3*.

Fig. 8 shows an example of the outputs of the methods under consideration for a patch with a strong ring artifact. The proposed 2D and 3D neural networks provide a very good visual quality of the corrected image. The result of the WFF method showed the good visual quality in the artifact area, but we can see strong blur the edges of objects outside the artifact area. Also, for processing by the WFF, a whole slice is required, while the proposed algorithms can process only the shown fragment.

The processing time for the fragment with a size of 512×512 voxels is about 0.025 seconds by the 2D network, and the fragment with $512 \times 512 \times 8$ voxels is about 0.377 seconds
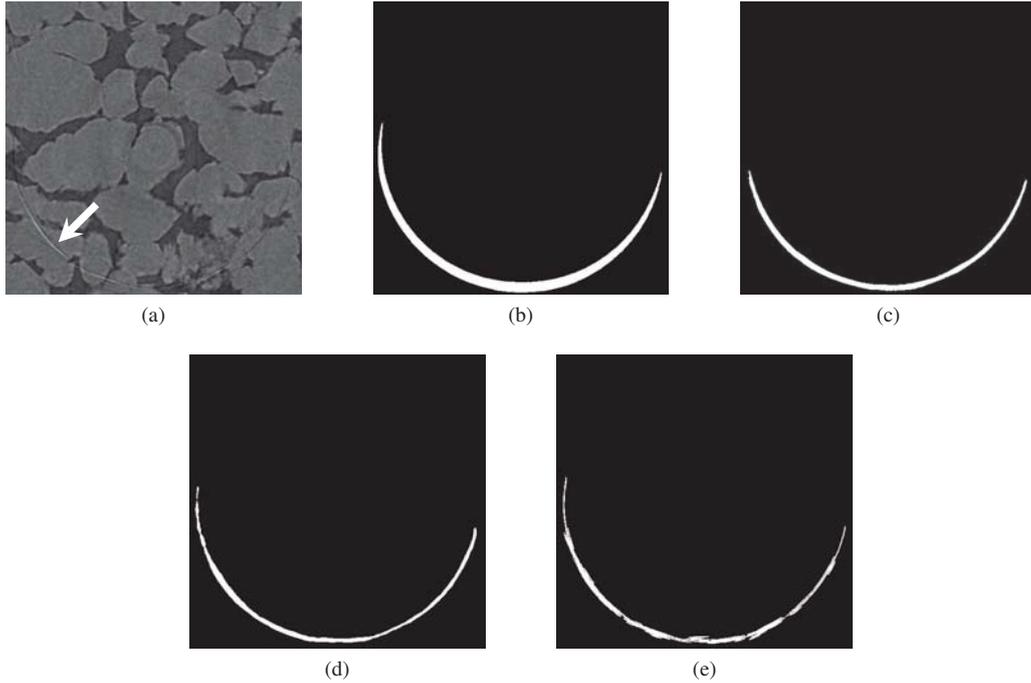
Fig. 6. Example of the ring artifact segmentation: (a) a fragment with ring artifact from the BHI_1 image; (b) GT; (c) mask obtained from the 2D CNN trained on the *Dataset 2*; (d) mask obtained from the 3D CNN trained on the *Dataset 2*; (e) mask obtained by [9]

TABLE III. THE RESULTS OF THE SEGMENTATION

Dataset	Dataset 1				Dataset 2			
Test set	UFS		UFS_r		BHI_1		$BHI_{1,r}$	
Metrics	Dice	PR-AUC	Dice	PR-AUC	Dice	PR-AUC	Dice	PR-AUC
Segmentation by [9]	0.70	0.73	0.35	0.53	0.77	0.80	0.43	0.59
Proposed 2D CNN	0.46	0.61	0.71	0.81	0.47	0.62	0.84	0.89
Proposed 3D CNN	0.56	0.72	**0.78**	**0.83**	0.55	0.76	**0.86**	**0.91**

TABLE IV. THE RESULTS OF THE RING ARTIFACTS CORRECTION

Dataset	Dataset 3			Dataset 4		
Metrics	MS-SSIM	MAE_{out}	MAE_{in}	MS-SSIM	MAE_{out}	MAE_{in}
No correction	0.9914	0.14	19.89	0.9858	0.10	12.75
Telea 1 [42] (masks from GT)	0.9944	0.14	14.28	0.9900	0.10	8.42
Telea 2 [42] (masks from 3D CNN)	0.9942	0.12	12.75	0.9886	0.12	8.42
WFF [10]	0.9823	2.11	10.97	0.9820	1.22	7.65
Partial convolution U-net [29]	0.9620	1.86	19.38	0.9542	1.01	10.97
Proposed 2D CNN	0.9974	0.31	7.65	0.9973	0.30	5.87
Proposed 3D CNN	0.9987	0.13	**4.76**	0.9992	0.09	**2.44**

by the 3D network on our GPUs using Pytorch with compute unified device architecture (CUDA).

VI. CONCLUSION

We proposed the 2D and 3D deep neural networks for the segmentation and correction of ring artifacts in microCT images of rock samples. We focus on artifacts remaining after conventional ring artifacts reduction methods in microCT system. Unlike existing approaches, the proposed method can be applied for fragments of a reconstructed CT image in the case

absence of shadow projection and the impossibility to make the conversion to polar coordinates. Also, our method does not require parameters tuning at the inference stage. The proposed 3D neural network outperforms methods under consideration for both segmentation and correction. The proposed 2D model provides quite high outcomes as well and operates faster in comparison with the 3D one. An application of the 2D network is a good trade-off between the quality, speed, and resources required for training.

One of the main contributions of our work is the method
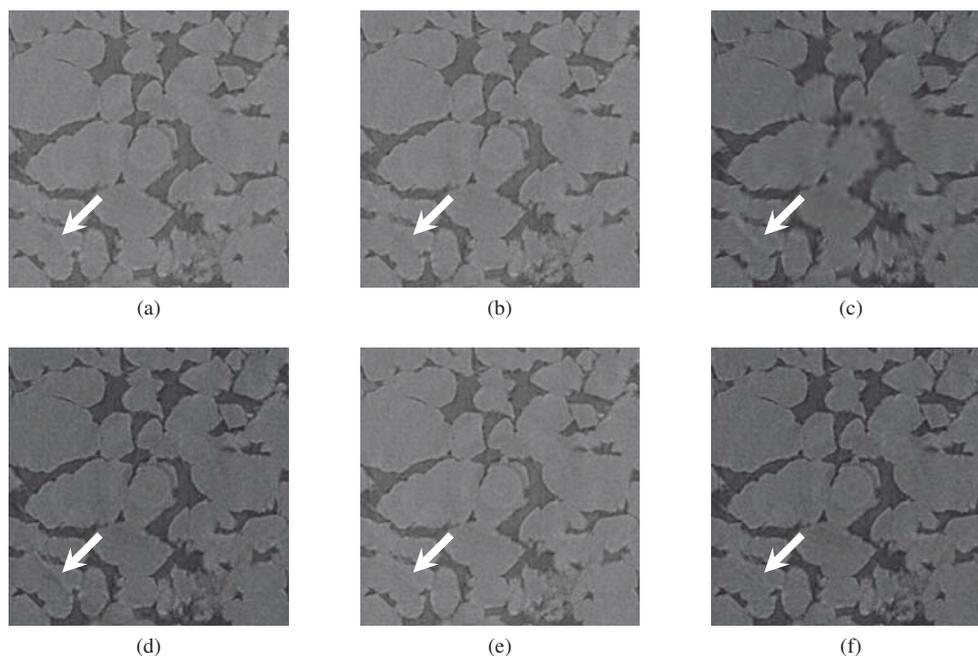
Fig. 7. Example of the ring artifact reduction from image in Fig. 6a: (**a**) with Telea's technique based on the mask from Fig. 6c; (**b**) with Telea's technique based on the mask from Fig. 6e; (**c**) with the WFF method; (**d**) with the U-net based on partial convolutions based on the mask from Fig. 6b; (**e**) with the 2D CNN trained on the *Dataset 4*; (**f**) with the 3D CNN trained on the *Dataset 4*

for the creation of training and testing datasets by transferring natural artifacts from one image to another. Our technique allows to generate realistic artifacts having various appearance due to the fusion of 3D ring and image in a random position of the cylindrical coordinate system. This method can be considered as specific data augmentation. Theoretically, there is a risk that the trained model will poorly detect and suppress artifacts of microCT systems from other manufacturers since they can differ significantly. However, the proposed algorithm for the translation of ring artifacts allows to rapidly create a dataset with many target artifacts and re-train the model. In addition, if an image contains a high level of noise, there is a high risk that the proposed approach will not work well.

In the future, we have the intention to investigate CycleGan for unpaired image-to-image translation [44] for the generation and suppression of ring artifacts. To reduce the time and memory costs during training for unpaired image-to-image translation, we are going to apply the Contrastive Unpaired Translation (CUT) approach, which is based on patch-wise contrastive learning and adversarial learning [45]. Also, we plan to extend the proposed approach for the inpainting of artifacts originated by high-density inclusions [5] and filling of cracks in the images of rock samples. Despite the proposed networks are intended for Digital Rock applications, our approach can be adjusted for the processing of CT images in other areas.

REFERENCES

[1] I. Safonov, I. Yakimchuk, and V. Abashkin, "Algorithms for 3D particles characterization using X-ray microtomography in proppant crush test," *Journal of Imaging*, vol. 4, no. 11, p. 134, 2018.

[2] I. Varfolomeev, I. Yakimchuk, and I. Safonov, "An application of deep neural networks for segmentation of microtomographic images of rock samples," *Computers*, vol. 8, no. 4, p. 72, 2019.

[3] A. Kornilov, I. Safonov, and I. Yakimchuk, "Inpainting of Ring Artifacts on Microtomographic Images by 3D CNN," in *2020 26th Conference of Open Innovations Association (FRUCT)*. IEEE, 2020, pp. 200–206.

[4] D. Koroteev, O. Dinariev, N. Evseev, D. Klemin, A. Nadeev, S. Safonov, O. Gurpinar, S. Berg, C. Van Kruijsdijk, R. Armstrong *et al.*, "Direct hydrodynamic simulation of multiphase flow in porous rock," *Petrophysics*, vol. 55, no. 04, pp. 294–303, 2014.

[5] J. F. Barrett and N. Keat, "Artifacts in CT: recognition and avoidance," *Radiographics*, vol. 24, no. 6, pp. 1679–1691, 2004.

[6] G. Davis and J. Elliott, "X-ray microtomography scanner using time-delay integration for elimination of ring artefacts in the reconstructed image," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 394, no. 1-2, pp. 157–162, 1997.

[7] M. Salehjahromi, Q. Wang, L. A. Gjesteby, D. Harrison, G. Wang, and H. Yu, "A directional TV based ring artifact reduction method," in *Medical Imaging 2019: Physics of Medical Imaging*, vol. 10948. International Society for Optics and Photonics, 2019, p. 109482C.

[8] W. Oh and B. Lindquist, "Image thresholding by indicator kriging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 7, pp. 590–602, 1999.

[9] A. Kornilov, I. Safonov, and I. Yakimchuk, "Blind Quality Assessment for Slice of Microtomographic Image," in *2019 24th Conference of Open Innovations Association (FRUCT)*. IEEE, 2019, pp. 170–178.

[10] B. Münch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet — Fourier filtering," *Opt. Express*, vol. 17, no. 10, pp. 8567–8591, May 2009.

[11] D. Gürsoy, F. De Carlo, X. Xiao, and C. Jacobsen, "TomoPy: a framework for the analysis of synchrotron tomographic data," *Journal of synchrotron radiation*, vol. 21, no. 5, pp. 1188–1193, 2014.

[12] I. Reimers, I. Safonov, and I. Yakimchuk, "Construction of 3D Digital Model of a Rock Sample Based on FIB-SEM Data," in *2019 24th Conference of Open Innovations Association (FRUCT)*. IEEE, 2019, pp. 351–359.

[13] D. Prell, Y. Kyriakou, and W. A. Kalender, "Comparison of ring artifact correction methods for flat-detector CT," *Physics in Medicine & Biology*, vol. 54, no. 12, p. 3881, 2009.
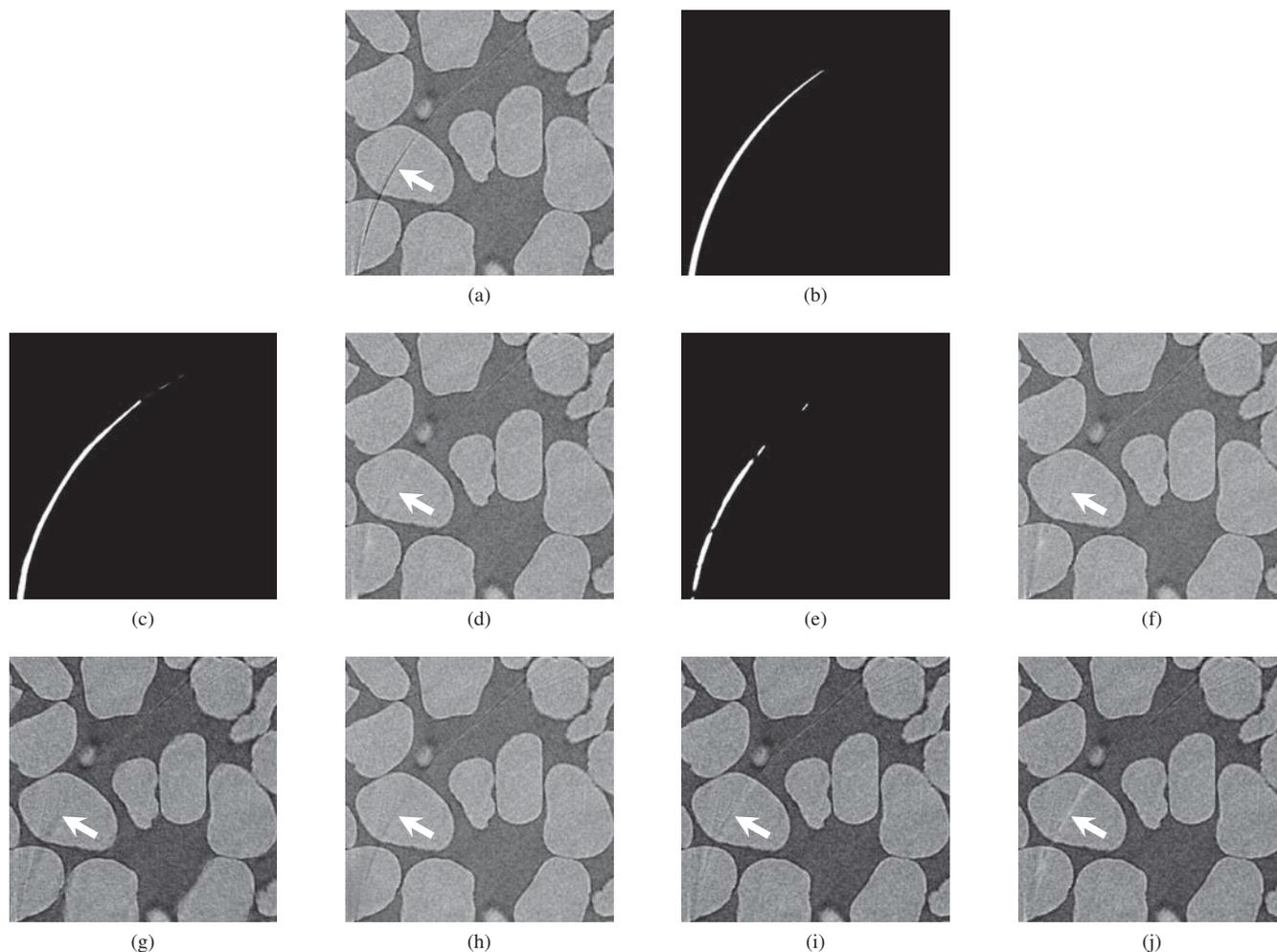
Fig. 8. An example of the strong ring artifact reduction: (a) fragment with ring artifact from the UFS image; (b) GT; (c) mask obtained from the 2D CNN trained on the *Dataset 1*; (d) corrected fragment with the Telea's technique based on the mask from Fig. 8c; (e) mask obtained from the 3D CNN trained on the *Dataset 1*; (f) corrected fragment with the Telea's technique based on the mask from Fig. 8e; (g) corrected fragment with the WFF method; (h) corrected fragment with the 2D CNN trained on the *Dataset 3*; (i) corrected fragment with the 3D CNN trained on the *Dataset 3*; (j) with the U-net based on partial convolutions based on the mask from Fig. 8b

[14] J. Sijbers and A. Postnov, "Reduction of ring artefacts in high resolution micro-CT reconstructions," *Physics in Medicine & Biology*, vol. 49, no. 14, p. N247, 2004.

[15] M. Axelsson, S. Svensson, and G. Borgefors, "Reduction of ring artifacts in high resolution X-ray microtomography images," in *Joint Pattern Recognition Symposium*. Springer, 2006, pp. 61–70.

[16] A. Lyckegaard, G. Johnson, and P. Tafforeau, "Correction of ring artifacts in X-ray tomographic images," *Int. J. Tomo. Stat*, vol. 18, pp. 1–9, 2011.

[17] E. M. A. Anas, J. G. Kim, S. Y. Lee, and M. K. Hasan, "Comparison of ring artifact removal methods using flat panel detector based CT images," *Biomedical engineering online*, vol. 10, no. 1, p. 72, 2011.

[18] S. Chang, X. Chen, J. Duan, and X. Mou, "A hybrid ring artifact reduction algorithm based on CNN in CT images," in *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, vol. 11072. International Society for Optics and Photonics, 2019, p. 1107226.

[19] W. Fang, L. Li, and Z. Chen, "Removing Ring Artefacts for Photon-Counting Detectors Using Neural Networks in Different Domains," *IEEE Access*, vol. 8, pp. 42 447–42 457, 2020.

[20] 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*. Springer, 2015, pp. 234–241.

[21] S. Xie, X. Zheng, Y. Chen, L. Xie, J. Liu, Y. Zhang, J. Yan, H. Zhu, and Y. Hu, "Artifact removal using improved GoogLeNet for sparse-view CT reconstruction," *Scientific reports*, vol. 8, no. 1, pp. 1–9, 2018.

[22] X. Huang, J. Wang, F. Tang, T. Zhong, and Y. Zhang, "Metal artifact reduction on cervical CT images by deep residual learning," *Biomedical engineering online*, vol. 17, no. 1, p. 175, 2018.

[23] W.-A. Lin, H. Liao, C. Peng, X. Sun, J. Zhang, J. Luo, R. Chellappa, and S. K. Zhou, "Dudonet: Dual domain network for ct metal artifact reduction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10 512–10 521.

[24] J. Xie, L. Xu, and E. Chen, "Image Denoising and Inpainting with Deep Neural Networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, ser. NIPS'12. Red Hook, NY, USA: Curran Associates Inc., 2012, p. 341–349.

[25] Y. Z. Ong, N. You, E. Li, and H. Yang, "Digital Rock Image Inpainting using GANs," 06 2020.

[26] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5505–5514.

[27] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and Locally Consistent Image Completion," *ACM Trans. Graph.*, vol. 36, no. 4, Jul.

2017.

[28] C. Yang, X. Lu, Z. Lin, E. Shechtman, O. Wang, and H. Li, "High-Resolution Image Inpainting Using Multi-Scale Neural Patch Synthesis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[29] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, "Image Inpainting for Irregular Holes Using Partial Convolutions," in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

[30] Z. Yan, X. Li, M. Li, W. Zuo, and S. Shan, "Shift-Net: Image Inpainting via Deep Feature Rearrangement," in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

[31] N. Abraham and N. M. Khan, "A novel focal tversky loss function with improved attention u-net for lesion segmentation," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 683–687.

[32] H. Kervadec, J. Bouchtiba, C. Desrosiers, E. Granger, J. Dolz, and I. B. Ayed, "Boundary loss for highly unbalanced segmentation," in *International conference on medical imaging with deep learning*, 2019, pp. 285–296.

[33] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 424–432.

[34] K. Lee, J. Zung, P. H. Li, V. Jain, and H. S. Seung, "Superhuman Accuracy on the SNEMI3D Connectomics Challenge," *arXiv*, vol. abs/1706.00120, 2017.

[35] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[36] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Transactions on computational imaging*, vol. 3, no. 1, pp. 47–57, 2016.

[37] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2. IEEE, 2003, pp. 1398–1402.

[38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[39] M. Kolařík, R. Burget, V. Uher, K. Říha, and M. Dutta, "Optimized High Resolution 3D Dense-U-Net Network for Brain and Spine Segmentation," *Applied Sciences*, vol. 9, no. 3, p. 404, Jan 2019.

[40] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PloS one*, vol. 10, no. 3, p. e0118432, 2015.

[41] S. Athar and Z. Wang, "A Comprehensive Performance Evaluation of Image Quality Assessment Algorithms," *IEEE Access*, vol. 7, pp. 140 030–140 070, 2019.

[42] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of graphics tools*, vol. 9, no. 1, pp. 23–34, 2004.

[43] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 2009, pp. 248–255.

[44] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017, pp. 2223–2232.

[45] T. Park, A. A. Efros, R. Zhang, and J.-Y. Zhu, "Contrastive Learning for Conditional Image Synthesis," in *ECCV*, 2020.