



Article Reprojection-Based Numerical Measure of Robustness for CT Reconstruction Neural Network Algorithms

Aleksandr Smolin^{1,2} ^(D), Andrei Yamaev ^{1,3,*(D)}, Anastasia Ingacheva ^{1,4} ^(D), Tatyana Shevtsova ⁵, Dmitry Polevoy ^{1,6,7}, Marina Chukalina ^{1,8}, Dmitry Nikolaev ^{1,4}, and Vladimir Arlazarov ^{1,6,*(D)}

- ¹ Smart Engines Service LLC, Moscow 121205, Russia
- ² Phystech School of Applied Mathematics and Informatics, Moscow Institute of Physics and Technology, Moscow 117303, Russia
- ³ Physics Department, Moscow State University, Moscow 119234, Russia
- ⁴ Institute for Information Transmission Problems (Kharkevich Institute) RAS, Moscow 127051, Russia
- ⁵ University Clinical Hospital No. 3 of the Clinical Center, I.M. Sechenov First Moscow State Medical
- University (Sechenov University), Moscow 119991, Russia
- ⁶ Federal Research Center Computer Science and Control RAS, Moscow 119333, Russia
- ⁷ National University of Science and Technology MISIS, Moscow 119049, Russia
- ⁸ FSRC Crystallography and Photonics RAS, Moscow 119333, Russia
- * Correspondence: rewin1996@gmail.com (A.Y.); vva@smartengines.com (V.A.)

Abstract: In computed tomography, state-of-the-art reconstruction is based on neural network (NN) algorithms. However, NN reconstruction algorithms can be not robust to small noise-like perturbations in the input signal. A not robust NN algorithm can produce inaccurate reconstruction with plausible artifacts that cannot be detected. Hence, the robustness of NN algorithms should be investigated and evaluated. There have been several attempts to construct the numerical metrics of the NN reconstruction algorithms' robustness. However, these metrics estimate only the probability of the easily distinguishable artifacts occurring in the reconstruction. However, these methods measure only the probability of appearance of easily distinguishable artifacts on the reconstruction, which cannot lead to misdiagnosis in clinical applications. In this work, we propose a new method for numerical estimation of the robustness of the NN reconstruction algorithms. This method is based on the probability evaluation for NN to form such selected additional structures during reconstruction which may lead to an incorrect diagnosis. The method outputs a numerical score value from 0 to 1 that can be used when benchmarking the robustness of different reconstruction algorithms. We employed the proposed method to perform a comparative study of seven reconstruction algorithms, including five NN-based and two classical. The ResUNet network had the best robustness score (0.65) among the investigated NN algorithms, but its robustness score is still lower than that of the classical algorithm SIRT (0.989). The investigated NN models demonstrated a wide range of robustness scores (0.38–0.65). Thus, in this work, robustness of 7 reconstruction algorithms was measured using the new proposed score and it was shown that some of the neural algorithms are not robust.

Keywords: robustness; neural network; computed tomography

MSC: 68T07

1. Introduction

One of the most popular methods of non-invasive investigation of the object's internal structure is X-ray computed tomography (CT) [1]. CT is widely used for pathology detection and treatment control in medicine [2–6], for flaw detection and quality control [7,8], or to solve different scientific problems, such as evaluation of the inner structure of pressurizers at nuclear power plants [9] or investigation of novel nanostructured materials [10].



Citation: Smolin, A.; Yamaev, A.; Ingacheva, A.; Shevtsova, T.; Polevoy, D.; Chukalina, M.; Nikolaev, D.; Arlazarov, V. Reprojection-Based Numerical Measure of Robustness for CT Reconstruction Neural Network Algorithms. *Mathematics* **2022**, *10*, 4210. https://doi.org/10.3390/ math10224210

Academic Editors: Hongang Qi, Yan Liu and Jun Miao

Received: 26 August 2022 Accepted: 4 November 2022 Published: 11 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). The internal structure of an object in CT is estimated from a set of measured projections, i.e., one-dimensional arrays of detector cell values taken at different angles and containing information about the absorption of X-rays passing through the object.

An important issue in medical CT is the reduction of radiation absorbed by a patient, because an excessive dose of X-ray radiation may lead to significant health problems for the patient. There are two general approaches to dose reduction: (1) reducing the exposure time, which results in noise within projections, and (2) reducing the number of measured projections. Currently, neural network (NN) algorithms demonstrate state-of-the-art reconstruction accuracy for low-angle and low-exposure tomography [11,12]. The applied NN reconstruction algorithms can be divided into three classes according to the application stage: preprocessing, postprocessing, and end-to-end algorithms. In preprocessing algorithms, neural networks are applied only to projections, for example, UNet1D [13]. In post-processing algorithms, neural networks are applied to the result of reconstruction obtained by analytical (Filtered Back Projection (FBP) [14]) or iterative algorithms (Simultaneous Iterative Reconstruction Technique (SIRT) [15]). FBPConvNet [16] and Residual U-Net (ResUNet) [17] are examples of the post-processing algorithms. End-to-end algorithms transform projections into a reconstruction by applying neural networks in both the projection space and the reconstruction space. This class of algorithms includes learned primal dual reconstruction (LPDR) [12] and TiraFL [18].

Inaccuracies or artifacts in reconstruction can lead to misdiagnosis. Classical algorithms (FBP, SIRT) produce well-known reconstruction artifacts [19]. These artifacts include metal-like [20], limited angle [21] artifacts, etc. Since the morphology of such artifacts is well known, their detection by specialists is not a problem. Thus, such artifacts may require only additional measurements in most cases. On the other hand, NN algorithms produce reconstruction artifacts that can be indistinguishable when the results are compared to the correct reconstructions [22]. The general scheme of difference between robust and not robust algorithms is demonstrated in Figure 1. Therefore, such artifacts may not be detected by a specialist. Thus, NN non-robustness can lead to misdiagnosis when using CT reconstruction exclusively. Such artifacts were shown in [22], where the NN algorithm [16] did not restore the reconstruction details characteristic of cancerous tumors by introducing insignificant noise to projections.



Figure 1. Demonstration of the difference between robust and not robust reconstruction algorithms.

The problem of the NN algorithms' robustness has been investigated previously. In [23], the authors showed the non-robustness of an NN image classification algorithm by introducing small perturbations into the input data. In [24], a similar non-robustness of the NN-based segmentation algorithm was shown for high-resolution real-world images. In [25], the non-robustness of NN-based segmentation algorithms was shown for CT reconstruction images. While there are many works on the robustness of reconstruction algorithms [18,26–28], the authors of these papers came to different conclusions. In [18,26], it was stated that NN algorithms are as robust as analytical reconstruction algorithms. In [27,28], it was stated that NN algorithms are not robust. Thus, the problem of the robustness of reconstruction algorithms is not solved, and the question of robustness assessment remains open. In [22], the authors showed only the fact that the NN reconstruction algorithm for X-ray computed tomography is not robust to tiny noise. In [26], the robustness of NN algorithms to dataset changes, resistance to small projection perturbations, and resistance to preserving small image details were investigated. In [18,26,28], a metric for assessing the robustness of magnetic resonance imaging (MRI) NN models to small projection perturbations was proposed: the maximum decrease in the reconstruction accuracy for a fixed norm of the input data perturbations. However, such a metric does not directly assess the possibility of misdiagnosis based on a set of reconstructions. In [18], reconstructions were shown as metric calculation results. However, these reconstructions contained only obvious reconstruction artifacts, which can lead only to re-measurement. Thus, the problem of numerical measure of the neural network CT reconstruction robustness in terms of diagnosis remains unsolved.

In this paper, we propose a novel task-oriented score for measuring the robustness of reconstruction algorithms. The method evaluates the possibility that in the reconstructed image, the algorithm would generate structures that may lead to an error in the diagnosis. Such structures in the reconstructed image will be called "dangerous" structures. Each dangerous structure is a reconstructed image to which a binary mask, or a so-called reconstruction perturbation, is added. The classical and neural network reconstruction algorithms were evaluated via the proposed method by averaging the resulting scores for all reconstructions from the Low Dose CT Grand Challenge dataset and four reconstruction perturbations. Four different reconstruction perturbations emulate real diseases. These perturbations were simulated based on a medical atlas of lung diseases.

The rest of the paper is organized as follows. Section 2.1 describes the proposed numerical measure of the NN models' robustness. Section 2.2 describes the training and test datasets for the NN reconstruction algorithms. Section 2.3 describes selected structures for robustness assessment by the proposed method. Structures were selected based on the medical atlas [29]. Section 2.4 shows the list of measured models. Section 3 describes and discusses the results of NN and classical reconstruction algorithms' robustness assessments.

2. Materials and Methods

2.1. Methodology for Investigating the Robustness of CT Reconstruction Neural Networks

In CT, the X-ray absorption by studied volume is measured with an X-ray detector. The measured values are calculated as follows:

$$I_{i} = I^{(0)} e^{-\int_{0}^{1} R(\vec{s}_{i} + (\vec{d}_{i} - \vec{s}_{i})t) \|\vec{d}_{i} - \vec{s}_{i}\| dt},$$
(1)

where $I^{(0)}$ is the number of initial photons directed at the detector cell ($I^{(0)}$ is the same for every cell), $\vec{d_i}$ is a detector position during the *i*th measurement, $\vec{s_i}$ is the X-ray source position during the *i*th measurement, $R(\vec{r})$ is a linear attenuation coefficient distribution within the volume under study, and I_i is the number of photons measured by the detector. Thus, the integral of $R(\vec{r})$ along the ray from $\vec{s_i}$ to $\vec{d_i}$ can be calculated reducing to the linear form the Equation (1)

$$p_i = -\ln \frac{I'_i}{I^{(0)}} = \int_0^1 R(\vec{s}_i + (\vec{d}_i - \vec{s}_i)t) \|\vec{d}_i - \vec{s}_i\| dt.$$
(2)

The set of p_i values obtained with the measurements I_i are called a "sinogram". $R(\vec{r})$ is a two-dimensional attenuation coefficient distribution. Let us approximate $R(\vec{r})$ as a two-

dimensional image R_{xy} with X rows and Y columns. Index j = yX + x; this is responsible for the image pixel position. Then values of p_i can be calculated as follows:

$$p_i = \sum_{j=1}^{n} a_{ij} R_j, \tag{3}$$

where $\{a_{ij}\}$ is an element of weight matrix, which represents a relationship between $\{R_{xy}\}$ and p_i . Value of a_{ij} equal to the length of the intersection between image pixel *j* and X-ray *i* as shown in Figure 2.



Figure 2. An example of I_i X-ray propagation along the pixels R_j of a reconstructed image in a parallel scheme.

Formula (3) is a direct representation of matrix multiplication:

$$P = AR_0, (4)$$

where *P* is a vector of all measured values, R_0 is a vector of all pixel values within the R_{xy} image, and *A* is a weight matrix. To calculate R_0 from (4), we will solve the inverse linear problem:

$$R = M(P), (5)$$

where $M(\cdot)$ is an arbitrary reconstruction algorithm. Let there be some reconstruction perturbation ΔR_0 that leads to an incorrect diagnosis. Then sinogram P' from the perturbed reconstruction R'_0 equals

$$R_0' = R_0 + \Delta R_0, \tag{6}$$

where R'_0 is a perturbed reconstruction.

$$P' = AR'_0 = A(R_0 + \Delta R_0) + \nu = AR_0 + A\Delta R_0 + \nu.$$
(7)

The sinogram perturbation ΔP of the original sinogram can be estimated as

L

$$\Delta P = P' - P, \tag{8}$$

$$\Delta P = A \Delta R_0, \tag{9}$$

If a reconstruction algorithm can change the diagnosis under projection and the change is much smaller than ΔP , then such an algorithm is not robust. The minimum difference between projections leading to a change in diagnosis can be calculated by solving the optimization problem:

$$\Delta P_M = \arg\min_{\Delta P'_M} \|M(P) + \Delta R_0 - M(P + \Delta P'_M)\|^2 + \lambda \|\Delta P'_M\|^2,$$
(10)

where ΔP_M is the perturbation with minimal norm which is required for the chosen perturbation ΔR_0 , $M(\cdot)$ is the reconstruction algorithm, and λ is the regularization parameter. From now on, we will refer to $||M(P) + \Delta R_0 - M(P + \Delta P'_M)||^2$ as "target reconstruction error rate".

Thus, we will calculate the robustness score as the difference in the norms of projections' deviations

$$S_M = 1 - \left| \frac{\|\Delta P\|^2 - \|\Delta P_M\|^2}{\|\Delta P\|^2} \right|.$$
(11)

In a robust reconstruction algorithm, S_M is expected to be approximately equal to 1 for any ΔR_0 . In Appendix A, we provide the explicit form of S_M for any linear reconstruction algorithm. Moreover, we demonstrate that the proposed score is always equal to 1 for the ideal linear algorithm for the tomography problem under the noise-free conditions.

To find ΔP_M from minimization problem (10) in a general case, we employed the L-BFGS (Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm) [30] algorithm with $\lambda = 1$. L-BFGS is the optimization algorithm in a family of quasi-Newton methods. L-BFGS evaluates the inverse Hessian matrix, the inverse matrix of second derivatives. The peculiarity of this method is that it has the proven property of convergence in a finite number of iterations for a linear equation system. This makes it applicable to this class of problems of finding the minimum perturbation. We select $\lambda = 1$ for the following reasons. For $\lambda = 10$, regularization has a noticeable effect on ΔP_M , for $\lambda = 0.1$, we did not observe noticeable changes in S_M ; however, the number of L-BFGS algorithm iterations increased. Therefore, we consider $\lambda = 1$ to be a near-optimal value for S_M calculation.

2.2. Data and Modeling

The NN algorithms for CT reconstruction were trained and their robustness was evaluated for the open dataset of medical tomography data "Low Dose CT Grand Challenge" [31–33]. The dataset comprises 99 head scans, 100 chest scans, and 100 abdomen scans, which were measured by Siemens and General Eclectic CT-scanners. The pixel width and height of each slice is 512. The example of chest reconstruction is shown in Figure 3.



Figure 3. Chest reconstruction from the "Low Dose CT Grand Challenge" dataset.

The dataset also contains the measurements used for reconstructed images. Projection data are provided within the dataset acquired by medical CT scanners in a helical scanning scheme (Figure 4A). However, most NN models for CT reconstruction have been developed to work with data collected based on a circular scanning scheme (Figure 4B). Therefore, for training and testing NN models, we used the projections which were simulated as follows.



Figure 4. Images of two measurement schemes: (A) helical scheme, (B) circular scheme.

We simulated the projections under a two-dimensional parallel-beam measurement scheme. This scheme suggests that the X-ray detectors are arranged along a line as illustrated in Figure 2. The rays directed at these cells are parallel to one another. We simulated 512 detectors, and the dimensions of each corresponded to that of the pixel within the considered region. According to Figure 4B, the measured volume is additionally rotated around its axis by a fixed step of $\frac{\pi}{512}$. We simulated 512 rotation steps to uniformly cover the range between 0 and 180 degrees. Therefore, for each reconstructed image, we simulated 512 × 512 projection values. The noise was simulated according to the Poisson distribution as follows.

$$I'_{i} = Poisson(I) = Poisson(I^{(0)}e^{-p_{i}}).$$
(12)

On the other hand, I'_i can be represented as

$$I'_i = I^{(0)} e^{-p'_i}. (13)$$

So, we can calculate noisy p'_i

$$p'_{i} = -\ln \frac{Poisson(I^{(0)}e^{-p_{i}})}{I^{(0)}},$$
(14)

where p'_i is the noisy sinogram value, Poisson(x) is a realization of a random variable: the Poisson distribution with x mean and dispersion of the distribution. This noise operation is applied to all values of all calculated projections with $I^{(0)} = 10^5$. This number of photons roughly corresponds to the number of photons received by X-ray detectors in medical CT-scanners. Each projection contains 512 values. The original high-dose reconstructions and noisy sinograms were used to train and test NN algorithms.

2.3. Reconstruction Perturbations

Based on the medical atlas [29], four reconstruction perturbations (pathologies) for the reconstructed images were formed. The perturbations corresponded to the airway involution and obstruction. These four perturbations were added to reconstructed images to construct ground truth for the L-BFGS algorithm. In (10), these perturbations are ΔR_0 . Perturbations and examples of their addition to the reconstructed image are shown in Figure 5. Simulated perturbations are binary masks without internal structure and these perturbations are placed in different parts of the lungs and have different shapes and sizes. These specific pathologies are relatively small and have simple inner structure; hence, they were chosen for the simulation. The localization of the simulated perturbations allows for a minimal distortion of the sinogram to ensure reproducibility. Images A and C show resuscitation and obstruction of the airways. This may be the sole sign of infection, typical for a viral one, such as in this example of a patient with a rhinovirus infection. Images B and D show asymmetric central bronchiectasis and mucoid occlusion, which is characteristic



of allergic bronchopulmonary aspergillosis and tuberculosis. When calculating our taskoriented robustness score (11), these reconstruction perturbations acted as ΔR_0 .

Figure 5. (A,C) Airway inflammation and airway impaction. (B,D) Asymmetric central bronchiectasis and mucoid impaction, which is typical for allergic bronchopulmonary aspergillosis, and tuberculosis.

2.4. Neural Network Models

We investigate the issue of robustness of five NN algorithms: ResUNet [17], FBPConvNet [16], LPDR [12], TiraFL [18], Unet1D [13]. We chose these algorithms to uniformly cover all classes of neural network algorithms for CT reconstruction. UNet1D was chosen as a representative of the preprocessing reconstructive network class. ResUNet and FBPConvNet were chosen as representative of the postprocessing reconstructive network class. LPDR and TiraFL were chosen as representative of the end-to-end reconstructive network class.

The FBPConvNet algorithm relies on applying the FBP algorithm (Filtered Back Projection) [14] to the sinogram, and then the FBP result is used as input of the neural network. FBPConvNet was proposed for the small angle computed tomography problem. The used U-Net network allowed to decrease the effect of small angle reconstruction artifacts. These artifacts are commonly presented as lines with size equal to all reconstructed image sizes. U-Net is widely used neural network architecture for segmentation [34] and its modifications are also widely used for denoising [13,16,17]. The algorithm uses a neural network with a modified version of U-Net architecture [34]. Unlike the case of U-Net, the output of this neural network algorithm is added to the input data. Thus, the neural network learns to predict the difference between the ideal reconstructed image and the reconstruction obtained using the FBP algorithm.

The ResUNet algorithm is similar to the FBPConvNet. The ResUNet algorithm relies on applying the FBP to the sinogram, and the result is used as input of the U-Net neural network. Unlike FBPConvNet, in ResUnet there are replaced "max pool" layers to "average pool" layers. This allows one to increase the robustness of a neural network to noise on the input data.

LPDR modifies the iterative SIRT (Simultaneous Iterative Reconstruction Technique) algorithm. In the SIRT algorithm, the reconstructed image is changed step by step as follows: (1) the sinogram from the current reconstructed image is calculated; (2) the difference between the measured sinogram and the computed sinogram is calculated; (3) this difference is backprojected into the reconstruction space and (4) is added to the current reconstructed image. In the LPDR algorithm, steps (2) and (4) are replaced by neural networks. In [12], a set of small convolutional neural networks is used. Each such NN included three consecutive 32 channel convolutions. LPDR contains only 10 modified iterations of the SIRT algorithm. LPDR is a fully-learned network that accepts sinograms as input and outputs a reconstructed image.

TiraFL is another fully-learned network. It was proposed to replace the back-projection linear operator with a trainable linear operator which is applied to the values of the sinogram, and the result is processed by a neural network based on the Tiramisu architecture. The trainable linear operator is a matrix with values that are calculated during the neural network learning process. It allows one to optimise the present relationship between reconstruction and projection value spaces.

The UNet1D algorithm is a modification of the FBP algorithm. The FBP reconstruction algorithm includes two stages: filtering one-dimensional projections with a ramp filter, and back-projecting the filtering results to obtain a reconstructed image. In the UNet 1D algorithm, the projection filtering stage is replaced by a one-dimensional neural network, which is a one-dimensional modification of the U-net neural network. According to the authors, this approach [13] guarantees a significant increase in the reconstruction accuracy without a large computational cost.

All neural networks were trained using the Adam optimizer [35] with an initial learning step of 1×10^{-4} and a batch size of 1. Moreover, all neural networks, with the exception of LPDR, were trained with a fixed learning step until convergence on the training set. In this paper, the neural networks with the best loss on the validation set were used. LPDR was trained over 100,000 learning steps and we used a cosine annealing [36] learning rate scheduler according to the original LPDR article. The MSE (mean squared error) loss function was employed for training the neural networks. The training and validation datasets were generated by a random sampling of reconstructed images from the Low Dose CT Grand Challenge dataset.

We implemented all neural network reconstruction algorithms using PyTorch library [37] and Astra Toolbox [38] to make the proposed score calculation uniform.

3. Results

We tested the proposed task-oriented score only for reconstructed images of lungs because we introduced perturbations based on lung diseases. Additionally, all lung reconstructed images did not contain any natural structures in the areas of proposed masks from Section 2.3. For reconstruction perturbation A, shown in Figure 5, the task-oriented robustness scores of NN algorithms (11) were calculated. λ from Equation (10) was set to 1 in all experiments. The calculation results are shown in Table 1 after 300 steps of the L-BFGS optimization algorithm. Plots of the deviation norm $||M(P) + \Delta R_0 - M(P + \Delta P'_M)||^2$ from (10) and the L2 norm $\Delta P'_M$ for the L-BFGS algorithm steps are shown in Figure 6. All variables have reached stable values. This means that 300 iterations of the L-BFGS optimization algorithm are enough to solve the given optimization problem. Each step of the L-BFGS algorithm requires calculation of a completely new reconstruction by the selected algorithms. In Figure 6, each iteration on the plot corresponds to a new reconstruction. Thus, 300 SIRT reconstructions are required to measure the proposed task-oriented score for the SIRT algorithm.

Table 1. Robustness scores calculated for "perturbation A" shown in Figure 5 over a single image (also shown in Figure 5).

Reconstruction Algorithm	Proposed Task-oriented Robustness Score	Score (15) of the Maximum Reconstruction Quality Degradation from a Slight Sinogram Perturbation
SIRT	0.96	20
FBP	0.87	74
SIRT-TV	0.82	18
ResUNet	0.64	30
TiraFL	0.60	39
UNet1D	0.52	33
LPDR	0.45	735
FBPConvNet	0.38	29



Figure 6. (**A**) Sinogram perturbation norms versus iteration number of the optimizing algorithm L-BFGS for various reconstruction algorithms. (**B**) Target reconstruction error rate from (10) versus the L-BFGS iteration number for various reconstruction algorithms. All plots are provided for the reconstructed images shown in Figure 7.

Additionally, we measured the robustness score already proposed in [26]. It is calculated as follows:

$$V = \frac{\min_{\Delta P: \|\Delta P\| \le \epsilon} - \frac{1}{2} \|M(P) - M(P + \Delta P)\|^2}{\epsilon},$$
(15)

where M(P) is a reconstruction algorithm, P is a sinogram, ΔP is a sinogram perturbation, ϵ is some fixed maximal norm of ΔP . We set ϵ to $\|\Delta P_0\|$ from Table 1 in order to ensure the correct comparison of network robustness scores. Moreover, we visualized $M(P + \Delta P)$ (see Figure 8) to demonstrate disadvantages of this score. The perturbed reconstructed images contain only visually detectable reconstruction artifacts that cannot lead to an inaccurate diagnosis based on a set of CT reconstructed images.

Tables 2 and 3 illustrate that FBP and SIRT algorithms are the most robust algorithms. On the other hand, ResUnet is the most robust neural reconstruction algorithm among all tested algorithms. For a visual assessment of the robustness for these algorithms, we calculated corresponding reconstructions (see Figure 7). Figure 7 is structured as follows. Each row shows reconstructions by the algorithm as indicated by the row label. The input of these reconstruction algorithms is the perturbed sinograms calculated by Formula (10) for algorithms which are indicated by the column label. Hence, we can observe the perturbations which are specific for certain reconstruction algorithms in terms of other

 Upper
 EPP
 Unicidi
 EPPC convoluti
 LPDR
 Tra
 SIRT
 SIRT

 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00
 00</td

reconstruction algorithms. The perturbed sinograms acquired by non-robust algorithms and used as input for other reconstruction algorithms produce reconstructions that are very different compared to the ideal one.

Figure 7. The figure shows areas of lung reconstructed images with an added artifact which simulates inflammation. Reconstructed images in each row of the table were created from one perturbed sinogram calculated for the neural network as indicated in the row label. Each column of the table contains reconstructed images created via an algorithm as indicated in the column labels.

The proposed task-oriented robustness score preserves the the algorithm ranking when its value is averaged over a set of reconstructed images (Table 2) and over a set of different masks (Table 3). The most robust algorithm is SIRT as illustrated in Table 3. The least robust algorithm is FBPConvNet. According to Tables 1–3, the rank of each algorithm is the same across the tables, except for the positions of the FBP and SIRT + TV algorithms. Throughout the calculations carried out for Table 3, we did not observe any significant changes in the values of the proposed metric for various masks.



Figure 8. Images with distorted reconstructions $M(P + \Delta P)$ from Equation (15). The title of each image corresponds to the reconstruction method.

Table 2. Averaged over 30 reconstructed images, the proposed task-oriented robustness score of the reconstruction algorithms calculated for the reconstruction perturbation A applied to the reconstructed image is shown in Figure 5.

Reconstruction Algorithm	Proposed Task-Oriented Robustness Score
SIRT	0.90
FBP	0.85
SIRT-TV	0.90
ResUNet	0.66
TiraFL	0.61
UNet1D	0.47
LPDR	0.46
FBPConvNet	0.40

Table 3. Proposed task-oriented score averaged over a set of 4 reconstruction perturbations and over 7 reconstructed images.

Reconstruction Algorithm	Proposed Task-Oriented Robustness Score
SIRT	0.989
FBP	0.997
SIRT-TV	0.892
ResUNet	0.649
TiraFL	0.594
UNet1D	0.486
LPDR	0.427
FBPConvNet	0.379

4. Discussion

The paper proposes a novel method for the numerical assessment of the neural network-based reconstruction algorithms' robustness. The measure is based on the reprojection of the selected reconstructed image perturbation and the minimum perturbation of the sinogram necessary for the appearance of the selected reconstructed image perturbation.

The proposed method evaluates the possibility of the appearance of selected perturbations in the reconstructed image. With the correct choice of perturbations, the proposed method evaluates the possibility of making an incorrect diagnosis when using the reconstructed image exclusively.

However, the proposed method also has some disadvantages. First, it depends on the chosen perturbations for the reconstructed image. Second, this paper does not mathematically prove that the reprojection from the reconstructed image based on synthetic low-exposure sinograms (from Section 2.2) is close in the ℓ_2 norm to the reconstructed image based on high-exposure sinograms. Although in certain cases the distribution function of values in the reconstructed image may assume the shape of a normal distribution, in general, the noise in the reconstructed image is more complex and requires additional study.

Computed tomography allows for the early diagnosis of various diseases. During the examination, the patient receives a dose of X-rays. Decreasing the dose may lead to degraded reconstructed image quality. Artificial intelligence technologies make it possible to strike a balance between reconstruction quality and radiation dose. However, neural network methods require a detailed study of their robustness; this will be done in future works.

5. Conclusions

In this work, we consider the CT reconstruction NN algorithms' robustness in terms of prominent features emerging in the reconstructed image due to noise-like distortions in the input data. This problem is crucial for medical applications since such artifacts may lead to a misdiagnosis.

We proposed a novel task-oriented robustness score that numerically measures neural network algorithm robustness. This score is based on the ℓ_2 norm of reprojection of the reconstruction perturbation as well as on the norm of the minimum sinogram perturbation, which is sufficient to generate the doubtful structure. Emergence of such a structure within the reconstructed images results in misdiagnosis. We calculated the proposed task-oriented score for five NN reconstruction algorithms as well as for both SIRT and FBP. According to our experiments, ResUNet has shown the best robustness score (0.65, see Table 3) among the investigated NN algorithms. However, in other NN algorithms, LPDR (0.43, see Table 3) and FBPConvNet (0.38, see Table 3), even much smaller sinogram deviations than expected can change the diagnosis when it is based on the reconstruction. This shows that LPDR and FBPConvNet are not sufficiently robust reconstruction methods. We have shown that both algorithms can output additional structures in lung CT data.

Author Contributions: Conceptualization, D.N. and V.A.; Methodology, A.Y., D.N. and V.A.; Project administration, M.C.; Software, A.S.; Supervision, D.N. and V.A.; Validation, T.S.; Visualization, A.Y.; Writing—original draft, A.Y.; Writing—review & editing, A.S., A.I., D.P., M.C., D.N. and V.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by RFBR (grant No. 19-29-09075, grant No. 19-29-09092).

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data used in this paper are publicly available under "LDCTand-Projection-data" [31–33] at https://doi.org/10.7937/9npb-2637. We used the 2016 Lowdose CT Challenge training set (accessed on 20 August 2022).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Properties of the Proposed Task-Oriented Score for Linear Reconstruction Algorithm

1

Let us consider a linear reconstruction algorithm

$$R = BP, \tag{A1}$$

where *B* is a matrix which represents a reconstruction algorithm. Let us calculate ΔP_M as follows:

$$\Delta P_M = \arg\min_{\Delta P'_M} \|BP + \Delta R_0 - B(P + \Delta P'_M)\|^2 + \lambda \|\Delta P'_M\|^2, \tag{A2}$$

where ΔR_0 is a chosen arbitrary perturbation of reconstruction, λ is a regularization parameter. We can simplify equation by expanding the brackets

$$\Delta P_M = \arg\min_{\Delta P'_M} \|BP + \Delta R_0 - BP - B\Delta P'_M)\|^2 + \lambda \|\Delta P'_M\|^2, \tag{A3}$$

$$\Delta P_M = \arg\min_{\Delta P'_M} \|\Delta R_0 - B\Delta P'_M)\|^2 + \lambda \|\Delta P'_M\|^2.$$
(A4)

Equation (A4) is a well-known optimization problem with Tikhonov regularization [39], and the solution is as follows:

$$\Delta P_M = (B^T * B + \lambda I)^{-1} B^T \Delta R_0. \tag{A5}$$

Then, we can calculate the proposed task-oriented score S_M without the minimization problem:

$$S_M = 1 - \left| \frac{\|\Delta P\|^2 - \|(B^T * B + \lambda I)^{-1} B^T \Delta R_0\|^2}{\|\Delta P\|} \right|.$$
(A6)

We can introduce the known value of $\Delta P = A \Delta R_0$, where *A* is a projection matrix.

$$S_M = 1 - \left| \frac{\|A \Delta R_0\|^2 - \|(B^T * B + \lambda I)^{-1} B^T \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right|.$$
(A7)

Based on Equation (A7), we derived 5 properties of the proposed evaluation function:

- 1. the proposed task-oriented score for any linear algorithm depends only on the chosen perturbation ΔR_0 ;
- 2. for any linear algorithm, the value of the proposed score does not change if ΔR_0 is multiplied by any constant $\alpha \neq 0$;
- 3. if $\lambda = 0$ and $B = A^{-1}$, then $S_M = 1$;
- 4. $S_M = 1$ for $\forall B : AA^T TT^T = 0$, where $T = (B^T * B + \lambda I)^{-1}B^T$;
- 5. $S_M = 1$ for $\forall B : (B^T * B + \lambda I)^{-1} B^T = A$.

The proofs of properties 1 and 5 are trivial, so they will not be provided here.

The proof of property 2 can be introduced as follows. Let $S_M(\alpha)$ be the proposed taskoriented score, which was calculated for multiplied reconstruction perturbation $\alpha \Delta R_0$ then

$$S_M(\alpha) = 1 - \left| \frac{\|A(\alpha \Delta R_0)\|^2 - \|(B * B + \lambda I)^{-1} B(\alpha \Delta R_0)\|^2}{\|A(\alpha \Delta R_0)\|^2} \right|,$$
(A8)

$$S_M(\alpha) = 1 - \left|\frac{\alpha^2 \|A \Delta R_0\|^2 - \alpha^2 \|(B^T * B + \lambda I)^{-1} B^T \Delta R_0\|^2}{\alpha^2 \|A \Delta R_0\|^2}\right|,\tag{A9}$$

$$S_M(\alpha) = 1 - \frac{\alpha^2}{\alpha^2} \left| \frac{\|A \Delta R_0\|^2 - \|(B^T * B + \lambda I)^{-1} B^T \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right| = S_M,$$
(A10)

$$_{M}(\alpha) = S_{M}. \tag{A11}$$

The proof of property 3 can be achieved via a direct introduction of values $B = A^{-1}$ and $\lambda = 0$ in Equation (A7)

S

$$S_M = 1 - \left| \frac{\|A \Delta R_0\|^2 - \|(B^T * B + \lambda I)^{-1} B^T \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right|,$$
(A12)

$$S_M = 1 - \left| \frac{\|A \Delta R_0\|^2 - \|((A^{-1})^T * A^{-1})^{-1} (A^{-1})^T \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right|,$$
(A13)

$$S_M = 1 - \left| \frac{\|A \Delta R_0\|^2 - \|((A^T)^{-1} * A^{-1})^{-1} (A^T)^{-1} \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right|,$$
(A14)

$$S_M = 1 - \left| \frac{\|A \Delta R_0\|^2 - \|A A^T (A^T)^{-1} \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right|, \tag{A15}$$

$$S_M = 1 - \left| \frac{\|A \Delta R_0\|^2 - \|A \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right|, \tag{A16}$$

$$S_M = 1 - |\frac{0}{\|A\Delta R_0\|^2}|,$$
 (A17)

$$S_M = 1. \tag{A18}$$

Thus, the proposed robustness score for an ideal linear algorithm always equals 1 under the noise-free conditions and if there is no regularization.

The proof of property 4 can be achieved by representing the norm as the scalar product:

$$S_M = 1 - \left| \frac{\|A \Delta R_0\|^2 - \|(B^T * B + \lambda I)^{-1} B^T \Delta R_0\|^2}{\|A \Delta R_0\|^2} \right| = 1,$$
 (A19)

$$\|A\Delta R_0\|^2 - \|(B^T * B + \lambda I)^{-1} B^T \Delta R_0\|^2 = 0.$$
 (A20)
Let $T = (B^T * B + \lambda I)^{-1} B^T$ then

$$\|A\Lambda R_0\|^2 - \|T\Lambda R_0\|^2 = 0$$
 (A21)

$$\|A\Delta R_0\| = \|I\Delta R_0\| = 0, \tag{A21}$$

$$(A\Delta R_0, A\Delta R_0) - (T\Delta R_0, T\Delta R_0) = 0,$$
(A22)

$$(A\Delta R_0, A\Delta R_0) - (T\Delta R_0, T\Delta R_0) = 0,$$

$$(A22)$$

$$(AA^T \Delta R_0, \Delta R_0) - (TT^T \Delta R_0, \Delta R_0) = 0,$$

$$(A23)$$

$$(A24)$$

$$((AAT - TTT)\Delta R_0, \Delta R_0) = 0,$$
(A24)

$$AA^T - TT^T = 0, (A25)$$

Solving the Minimization Problem with Tikhonov Regularization

$$\Delta P_M = \arg\min_{\Delta P'_M} \|BP + \Delta R_0 - B(P + \Delta P'_M)\|^2 + \lambda \|\Delta P'_M\|^2, \tag{A26}$$

$$\Delta P_M = \arg\min_{\Delta P'_M} \|BP + \Delta R_0 - BP - B\Delta P'_M)\|^2 + \lambda \|\Delta P'_M\|^2, \tag{A27}$$

$$\Delta P_M = \arg\min_{\Delta P'_M} \|B\Delta P'_M - \Delta R_0\|^2 + \lambda \|\Delta P'_M\|^2.$$
(A28)

Let $K(\Delta P'_M) = \|B\Delta P'_M - \Delta R_0\|^2 + \lambda \|\Delta P'_M\|^2$ and then

$$\Delta P_M = \arg\min_{\Delta P'_M} K(\Delta P'_M). \tag{A29}$$

Equation (A29) can be solved through differentiation of $K(\Delta P'_M)$ by $\Delta P'_M$. The minimum ΔP_M must satisfy the condition

$$\frac{\partial K(\Delta P_M)}{\partial \Delta P_M} = 0. \tag{A30}$$

In accordance with the rules of differentiation of norms, the Equation (A30) can be expanded as follows:

$$\frac{\partial K(\Delta P_M))}{\partial \Delta P_M} = 2\|B\Delta P_M - \Delta R_0\|(\frac{B\Delta P_M - \Delta R_0}{\|B\Delta P'_M - \Delta R_0\|}, Bh) + 2\lambda\|\Delta P_M\|(\frac{\Delta P_M}{\|\Delta P_M\|}, h) = 0,$$
(A31)

where *h* is any vector. Next, we can simplify the equation as follows:

$$(B\Delta P_M - \Delta R_0, Bh) + \lambda(\Delta P_M, h) = 0, \tag{A32}$$

$$(B^{T}(B\Delta P_{M} - \Delta R_{0}) + \lambda \Delta P_{M}, h) = 0,$$
(A33)

$$B^{T}(B\Delta P'_{M} - \Delta R_{0}) + \lambda \Delta P_{M} = 0, \qquad (A34)$$

$$(B^T B + \lambda I)\Delta P_M = B^T \Delta R_0, \tag{A35}$$

$$\Delta P_M = (B^T B + \lambda I)^{-1} B^T \Delta R_0. \tag{A36}$$

Equation (A36) is a solving of minimization problem with Tikhonov regularization.

References

- Cherezov, D.; Paul, R.; Fetisov, N.; Gillies, R.J.; Schabath, M.B.; Goldgof, D.B.; Hall, L.O. Lung nodule sizes are encoded when scaling CT image for CNN's. *Tomography* 2020, *6*, 209–215. [CrossRef] [PubMed]
- 2. Mandell, J.C.; Khurana, B. Clinical applications of a computed tomography color "marrow mapping" algorithm to increase conspicuity of nondisplaced trabecular fractures. *Emerg. Radiol.* **2019**, *26*, 59–66. [CrossRef] [PubMed]
- 3. Baldacci, F.; Laurent, F.; Berger, P.; Dournes, G. 3D human airway segmentation from high hesolution MR imaging. *SPIE* **2019**, 11041, 244–250.
- Sharma, M.; Bhatt, J.S.; Joshi, M.V. Early detection of lung cancer from CT images: Nodule segmentation and classification using deep learning. SPIE 2018, 10696, 226–233.
- Pashina, T.A.; Gaidel, A.V.; Zelter, P.M.; Kapishnikov, A.V.; Nikonorov, A.V. Automatic highlighting of the region of interest in computed tomography images of the lungs. *Comput. Opt.* 2020, 44, 74–81. [CrossRef]
- Manzo, M.; Pellino, S. Fighting together against the pandemic: Learning multiple models on tomography images for COVID-19 diagnosis. AI 2021, 2, 261–273. [CrossRef]
- 7. Kirsch, S. Computed Tomography as a Tool for Archiving Ethnomusicological Objects. In *Computational Phonogram Archiving*; Springer: Cham, Switzerland, 2019; pp. 305–319.
- 8. Ho, H.; Yu, H.; Gangsei, L.; Kongsro, J. A CT-image based pig atlas model and its potential applications in the meat industry. *Meat Sci.* **2019**, *148*, 1–4. [CrossRef]
- Gratton, R.; Wilson, J.; Skuse, K. Application of Computed Tomography for the Examination of Pressure Retaining Nuclear Plant Components. In *Pressure Vessels and Piping Conference*; American Society of Mechanical Engineers: New York, NY, USA, 2017; p. V005T10A001.
- Bulatov, K.; Chukalina, M.; Kutukova, K.; Kohan, V.; Ingacheva, A.; Buzmakov, A.; Arlazarov, V.V.; Zschech, E. Monitored tomographic reconstruction—An advanced tool to study the 3D morphology of nanomaterials. *Nanomaterials* 2021, 11, 2524. [CrossRef]
- Chen, H.; Zhang, Y.; Zhang, W.; Liao, P.; Li, K.; Zhou, J.; Wang, G. Low-dose CT denoising with convolutional neural network. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 143–146.
- 12. Adler, J.; Öktem, O. Learned primal-dual reconstruction. IEEE Trans. Med. Imaging 2018, 37, 1322–1332. [CrossRef]
- 13. Yamaev, A.; Chukalina, M.; Nikolaev, D.; Sheshkus, A.; Chulichkov, A. Lightweight denoising filtering neural network for FBP algorithm. *SPIE* **2021**, *11605*, 158–167.
- Pan, X.; Sidky, E.Y.; Vannier, M. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Probl.* 2009, 25, 123009. [CrossRef] [PubMed]
- 15. Trampert, J.; Leveque, J.J. Simultaneous Iterative Reconstruction Technique: Physical interpretation based on the generalized least squares solution. *J. Geophys. Res. Solid Earth* **1990**, *95*, 12553–12559. [CrossRef]
- Jin, K.H.; McCann, M.T.; Froustey, E.; Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* 2017, 26, 4509–4522. [CrossRef] [PubMed]
- 17. Heinrich, M.P.; Stille, M.; Buzug, T.M. Residual U-Net convolutional neural network architecture for low-dose CT denoising. *Curr. Dir. Biomed. Eng.* **2018**, *4*, 297–300. [CrossRef]
- 18. Genzel, M.; Macdonald, J.; Marz, M. Solving inverse problems with deep neural networks-robustness included. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**. [CrossRef] [PubMed]

- Nikolaev, D.; Buzmakov, A.; Chukalina, M.; Ivan, Y.; Gladkov, A.; Ingacheva, A. CT image quality assessment based on morphometric analysis of artifacts. SPIE 2017, 10253, 52–56.
- Bal, M.; Spies, L. Metal artifact reduction in CT using tissue-class modeling and adaptive prefiltering. *Med. Phys.* 2006, 33, 2852–2859. [CrossRef]
- 21. Yu, W.; Zeng, L. 10 gradient minimization based image reconstruction for limited-angle computed tomography. *PLoS ONE* 2015, 10, e0130793. [CrossRef]
- 22. Huang, Y.; Würfl, T.; Breininger, K.; Liu, L.; Lauritsch, G.; Maier, A. Some investigations on robustness of deep learning in limited angle tomography. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2018; pp. 145–153.
- Maron, R.C.; Schlager, J.G.; Haggenmüller, S.; von Kalle, C.; Utikal, J.S.; Meier, F.; Gellrich, F.F.; Hobelsberger, S.; Hauschild, A.; French, L.; et al. A benchmark for neural network robustness in skin cancer classification. *Eur. J. Cancer* 2021, 155, 191–199. [CrossRef]
- 24. Arnab, A.; Miksik, O.; Torr, P.H. *On* the Robustness of Semantic Segmentation Models to Adversarial Attacks. *arXiv* 2018, arXiv:1711.09856.
- 25. Feng, X.; Bernard, M.E.; Hunter, T.; Chen, Q. Improving accuracy and robustness of deep convolutional neural network based thoracic OAR segmentation. *Phys. Med. Biol.* **2020**, *65*, 07NT01. [CrossRef] [PubMed]
- Darestani, M.Z.; Chaudhari, A.S.; Heckel, R. Measuring robustness in deep learning based compressive sensing. *PMLR* 2021, 139, 2433–2444.
- Zhang, C.; Jia, J.; Yaman, B.; Moeller, S.; Liu, S.; Hong, M.; Akçakaya, M. Instabilities in Conventional Multi-Coil MRI Reconstruction with Small Adversarial Perturbations. In Proceedings of the 2021 55th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 31 October–3 November 2021; pp. 895–899.
- Antun, V.; Renna, F.; Poon, C.; Adcock, B.; Hansen, A.C. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc. Natl. Acad. Sci. USA* 2020, *117*, 30088–30095. [CrossRef] [PubMed]
- 29. Elicker, B.; Webb, W. Fundamentals of High-resolution Lung CT: Common Findings, Common Patterns, Common Diseases, and Differential Diagnosis. *Pneumologie* **2013**, *67*, 400.
- 30. Liu, D.C.; Nocedal, J. On the limited memory BFGS method for large scale optimization. *Math. Program.* **1989**, 45, 503–528. [CrossRef]
- 31. Moen, T.R.; Chen, B.; Holmes, D.R., III; Duan, X.; Yu, Z.; Yu, L.; Leng, S.; Fletcher, J.G.; McCollough, C.H. Low-dose CT image and projection dataset. *Med. Phys.* 2021, *48*, 902–911. [CrossRef]
- Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. J. Digit. Imaging 2013, 26, 1045–1057. [CrossRef]
- McCollough, C.; Chen, B.; Holmes, D.; Duan, X.; Yu, Z.; Yu, L.; Leng, S.; Fletcher, J. Low Dose CT Image and Projection Data (LDCT-and-Projection-Data) (Version 4) [Data Set]. 2020. Available online: https://wiki.cancerimagingarchive.net/pages/ viewpage.action?pageId=5275802668b4ee803a824d47bfadb5a5f9563011 (accessed on 25 August 2022).
- Du, G.; Cao, X.; Liang, J.; Chen, X.; Zhan, Y. Medical image segmentation based on u-net: A review. J. Imaging Sci. Technol. 2020, 64, 1–12. [CrossRef]
- 35. Ruder, S. An overview of gradient descent optimization algorithms. *arXiv* **2016**, arXiv:1609.04747.
- 36. Loshchilov, I.; Hutter, F. Sgdr: Stochastic gradient descent with warm restarts. arXiv 2016, arXiv:1608.03983.
- 37. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* 2019, 32. Available online: https: //proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html (accessed on 25 August 2022).
- Van Aarle, W.; Palenstijn, W.J.; Cant, J.; Janssens, E.; Bleichrodt, F.; Dabravolski, A.; De Beenhouwer, J.; Batenburg, K.J.; Sijbers, J. Fast and flexible X-ray tomography using the ASTRA toolbox. *Opt. Express* 2016, 24, 25129–25147. [CrossRef] [PubMed]
- Golub, G.H.; Hansen, P.C.; O'Leary, D.P. Tikhonov regularization and total least squares. SIAM J. Matrix Anal. Appl. 1999, 21, 185–194. [CrossRef]