RISING: A new framework for model-based few-view CT image reconstruction with deep learning

RISING: a new framework for model-based few-view CT image reconstruction with deep learning

Davide Evangelista[a], Elena Morotti[b,*], Elena Loli Piccolomini[c]

[a] *Department of Mathematics, University of Bologna, Italy.*
[b] *Department of Political and Social Sciences, University of Bologna, Italy.*
[c] *Department of Computer Science and Engineering, University of Bologna, Italy.*

Abstract

Medical image reconstruction from low-dose tomographic data is an active research field, recently revolutionized by the advent of deep learning. In fact, deep learning typically yields superior results than classical optimization approaches, but unstable results have been reported both numerically and theoretically in the literature. This paper proposes RISING, a new framework for sparse-view tomographic image reconstruction combining an early-stopped Rapid Iterative Solver with a subsequent Iteration Network-based Gaining step. In our two-step approach, the first phase executes very few iterations of a regularized model-based algorithm, whereas the second step completes the missing iterations by means of a convolutional neural network.
The proposed method is ground-truth free; it exploits the computational speed and flexibility of a data-driven approach, but it also imposes sparsity constraints to the solution as in the model-based setting. Experiments performed both on a digitally created and on a real abdomen data set confirm the numerical and visual accuracy of the reconstructed RISING images in short computational times. These features make the framework promising to be used on real systems for clinical exams.

Keywords: Sparse tomography, Tomographic imaging, Deep Learning, Model-based iterative solver, Few-view tomography.

1. Introduction

Combining healthy protocols with high quality images is one of the most important component of medical imaging and a crucial target for researchers involved in minimal invasive Computed Tomography (CT). Radiologists, manufacturers and medical physicists have implemented many examination protocols

*Corresponding author
Email addresses: davide.evangelista5@unibo.it (Davide Evangelista), elena.morotti4@unibo.it (Elena Morotti), elena.loli@unibo.it (Elena Loli Piccolomini)

as well as software and hardware modifications to reduce the harmful ionizing radiations and pave the way to X-ray imaging for screening tests, pediatric cases, or pre-surgical exams. There are two main techniques allowing for a significant reduction of the total radiation exposure per patient. The first one consists in reducing the X-ray tube current at each scan (*low-dose CT*), without changing the traditionally used CT full geometry. The resulting measured data is very noisy, due to excessive quantum noise. The second practical way to lower the radiation per person consists in reducing the number of X-ray projections (*few-view CT*), which leads to incomplete tomographic data, but very fast examinations.

In this paper, we focus on few-view CT images, but we resonably argue that the proposed algorithm could be also applied with success to low-dose CT. In the case of sub-sampled data, the image reconstruction is tricky and conventional filtered backprojection (FBP) algorithms, widely exploited in classical CT, do not provide stable reconstructions as the recovered images typically suffer from severe striking artifacts. A widely used alternative approach is represented by model-based iterative methods. They model the image reconstruction as a mathematical linear inverse problem which is solved, in the discrete setting, by minimizing a constrained or unconstrained function combining a data-fitting term and a regularizer. Due to the lack of projections, the CT discrete inverse problem has infinite possible solutions [1]. The embedding of a sparsifying regularizer mitigates the lack of many projection views, according to the Compressed Sensing theory [2]. An exhaustive review of model-based reconstruction methods can be found in [3].
Recently, Deep Learning (DL) based methods have emerged over fully conventional and variational approaches for few-view tomographic reconstruction [4]. The two widely used strategies are the so called learnt model-based and the Learnt Post Processing (LPP) approaches. The learnt model-based schemes are iterative solvers (as classical optimization algorithms) where a neural network is used in each iteration to determine the best image update (i.e. the best parameters or operators defining it) [5, 6, 7]. On the contrary, the LPP is a two-step scheme where first a low quality image with artifacts and noise is reconstructed with a fast method (typically a FBP) and then a neural network suppresses the artifacts. Usually, the network learns from a set of *ground truth* images, reconstructed from full dose acquisitions in a long offline phase, whereas its online execution is characterized by a very fast performance. The pioneering works by Han in 2016 and 2018 demonstrated the superiority of LPP strategies over some model-based iterative algorithms for few-view CT images [8, 9].

Motivation. The main disadvantage of model-based iterative reconstruction algorithms is their high computational cost, since they typically need several iterations to achieve high quality results. On real systems, in order to fulfill the clinical requirements of time per exam, only very few iterations of the algorithms can be performed, leading to a solution which is not the best achievable one. However, by considering fast iterative methods, the images reconstructed in few iterations already contain many details of interest of the scanned object [10, 11].

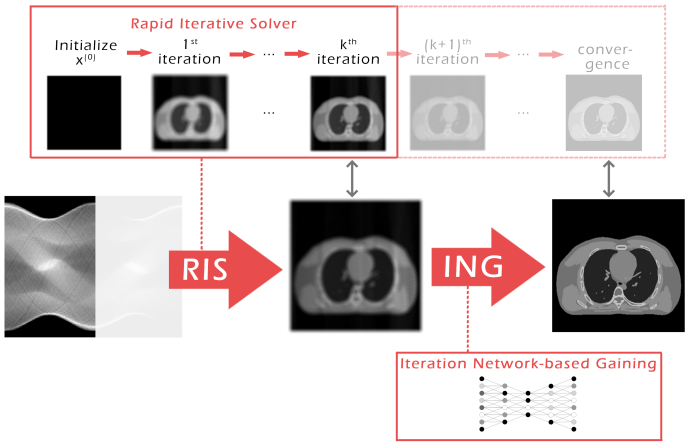

Figure 1: Graphical draft of the proposed two-step RISING workflow for tomographic reconstruction from few-view data.

Focusing on LPP based approaches, the work by Sidky et al. [12] claims and demonstrates that these schemes do not compute the solution of the CT inverse problem and they can introduce, in the reconstruction, structures not belonging to the scanned objects. On the contrary, numerical evidence shows that model-based methods compute a good solution of the inverse problem, according to the mathematical sense that will be defined in the next section. In addition, it is worth noticing that neural networks need to be trained on task-specific data sets to properly learn both the degradation effects to be removed and the anatomical details to be preserved, which characterize each typology of medical imaging; hence a further relevant challenge for medical applications is the lack of precise training data [4].

Proposal and contribution. Strong of this awareness, this paper proposes a new iterative and data-driven scheme, which does not require full-dose target images, for the resolution of the few-view CT inverse problem. We refer to our proposal as the *RISING* (Rapid Iterative Solver with Iteration Network-based Gaining) framework, whose graphical draft is depicted in Figure 1. RISING is conceived as one reconstructing procedure, described by the two following steps, executed in sequence:

- Starting from the sub-sampled projection data, a rapid model-based iterative algorithm produces a preliminary coarse reconstruction, in few iterations. The execution of only few iterations fits realistic time constraints.

- The computed rough reconstruction is post-processed by a pre-trained convolutional neural network, providing the RISING solution. By completing the iterations up to the iterative method convergence, we obtain images used as target for the neural network training. This ensures a fully consistent training set with respect to the system geometry, and it is

applicable to all types of medical images and few-view protocols.

This work has a dual purpose. On one side, we aim at combining the fast execution of LPP methods with the stability of model-based iterative algorithms used in sparse CT reconstruction. The variational robustness can face, in fact, the lack of projection data and the presence of noise on the projections, thanks to the use of suitable regularization functions acting as priors onto the solution. As the execution of a neural network is very fast, the use of DL greatly speeds up the whole reconstruction.
On the other side, we intend to remedy to the lack of task-specific CT data sets by using, for networks training, images created by the same system and under the same geometry used for the reconstruction. This is attainable for every CT system and makes the RISING framework suitable for real use.

To validate our proposal, we have applied the RISING approach on synthetic and real CT data sets simulating various few-view geometries. We developed our RISING workflow in Python and the code is available at `https://github.com/loibo/RISING`.

Organization of the paper. The paper is organized as follows. In Section 2 we present the state-of-art for the CT image reconstruction DL-based methods, then Section 3 describes the proposed RISING framework. In Section 4 we introduce the experimental settings considered to achieve the results reported and discussed in Section 5. At last, final conclusions are drawn in Section 6.

2. Deep methods for CT reconstruction

As anticipated, algorithms exploiting deep neural networks have been proposed for few-view CT image reconstruction for few years. Many publications have focused on learnt model-based approaches, where an iterative solver is unfolded into the sequence of its iterations [13] (the so called *unrolled* or *unfolded* approach). Even if they are thought as optimization methods applied to a general regularized physical model, since the parameters learnt by the network are different at each iteration, the final image is not the solution of the original minimization. Moreover, the neural networks employed for unfolded schemes suffer from instabilities, such as vanishing gradient [14]. To reduce these unwanted effects only a limited number of unfolded iterations is performed, forcing to stop the method far before convergence. The proposals in literature differ for the unrolled iterative scheme and for the block-per-iteration learnt by the neural network. We give a short overview of the most recent works. In [15], Adler and Oktem have developed a partially learnt gradient descent algorithm, whereas they have worked on the Chambolle-Pock scheme in [16]. In [17] a CNN is trained to act like a projector in a gradient descent algorithms. In [18] both the proximal operator and gradient operator (corresponding to A^T when it is applied to a problem corrupted with additive Gaussian noise) of an unrolled FISTA scheme are learnt, whereas in [19] the neural network learns the initial iterate of the inner Conjugate Gradients solver in a splitting scheme for optimization.

A different data-driven proposal for CT image reconstruction is represented by the LPP approach. It consists of a two-step framework where, firstly, a fast reconstruction algorithm is applied off-line to compute a coarse reconstruction of the medical image from the projection data, then a pre-trained neural network is applied as post-processing operator to restore the low-quality image. It is worth noting that most of the networks used for LPP are applied to corrupted images provided by the FBP solver (see for example [9, 20, 21, 22, 23, 24] and the references therein). Only in [25] and in [26] the preliminary solutions are computed by wavelet-based or TV-based iterative solvers, respectively.

3. The RISING framework

In this section we describe the proposed RISING framework by focusing on the Rapid Iterative Solver (RIS) phase in section 3.1 and on the Iteration Network-based Gaining (ING) phase in section 3.2. For simplicity, we consider in this paper only the two dimension case. The extension to three dimensions is straightforward.

3.1. Rapid Iterative Solver

In the discrete setting, the CT process of X-ray absorption is expressed as:

$$Ax = b \tag{1}$$

where the unknown vector $x \in \mathbb{R}^{n^2}$ denotes the image to reconstruct, the right-hand side term b is an m-dimensional vector containing the noisy projection measurements and the $m \times n^2$ system matrix A is the discretization of the X-ray physical process projecting an $n \times n$ image onto the detector. We say that a system of linear equations is *solvable* in a subset Ω if it admits a unique solution in Ω. In case of few-view CT protocols, Equation (1) is not solvable in $\mathbb{R}^{n^2}$, since $m < n^2$ and the under-determined linear system admits infinite solutions. According to the Compressed Sensing theory [2], if the desired solution x^* of (1) is sparse in some transform Tx, then Equation (1) is solvable in the subset $\Omega = \{x \in \mathbb{R}^{n^2} \mid ||Tx||_0 \leq d\}$, where $\|\cdot\|_0$ is the ℓ_0 semi-norm counting the non-zero elements of the vector argument. Hence the CT inverse problem can be reformulated as the following minimization:

$$x^* = \arg\min_x \|Tx\|_0 \text{ s.t. } Ax = b. \tag{2}$$

which admits a unique solution if A is full-rank.

Equation (2) is usually relaxed in the following, more easily solvable, form:

$$\arg\min_x \|Tx\|_1 \text{ s.t. } Ax = b, \tag{3}$$

where $\|\cdot\|_1$ is the ℓ_1 norm [27]. An unconstrained formulation of (3) can be stated as:

$$\arg\min_x \|Ax - b\|_2^2 + \lambda\|Tx\|_1 \tag{4}$$

where λ is a suitable positive parameter (also called *regularization parameter*). Convergent iterative schemes, such as the lagged diffusivity fixed point algorithm [28] or FISTA [29], are commonly exploited to solve (4). Sometimes, the constraint $x \geq 0$ is added to (3) or (4), to preserve the physical non-negativity property of the attenuation coefficients [30, 31, 11].

In the following we will consider the particular minimization problem:

$$\arg\min_{x \geq 0} \|Ax - b\|_2^2 + \lambda TV_\beta(x). \tag{5}$$

where TV_β is a a smoothed differentiable version of the Total Variation [32], obtained by setting T as the image gradient discretization, to exploit sparsity in the image gradient domain. It is defined as:

$$TV_\beta(x) = \sum_{j=1}^{n^2} \sqrt{\|\nabla x_j\|_2^2 + \beta^2} \tag{6}$$

with β fixed small positive parameter.

Among the wide class of iterative solvers for the problem (5), we select the Scaled Gradient Projection (SGP) algorithm, proposed in 2008 for image deblurring [33]. The SGP has been successfully applied with some acceleration techniques for few-view CT reconstructions in the papers [34, 10, 11], showing that the objects of interest are already distinguishable in the image computed after very few iterations of the method. Regarding the convergence, it is proved in [35] that the theoretical convergence rate of the SGP at the unique minimum of (5) is $\mathcal{O}(1/\mathrm{k})$, even if empirical tests reveal that the SGP performances are closely comparable to the convergence rate of optimal algorithms. For a detailed presentation of SGP method in tomographic imaging see [34].

The first step of the RISING framework performs a predefined (and relatively small) number K of iterations of the considered Rapid Iterative Solver (RIS). We remark that we stop the SGP far before its convergence to meet the constraints imposed by clinical setting, where very short computational times are admitted for the reconstruction process. The early solution $x^{(K)}$ is denoted as x_{RIS} in the following and it represents the input of the neural network which enhances the image in a new perspective.

3.2. Iteration Network-based Gaining

The second step of the RISING framework implements the Iteration Network-based Gaining (ING) task. From our previous works [34, 10] we know that, due to the empirical rapidity of the SGP, in the achieved coarse image reconstruction the anatomical structures are present but typically more blurred and less visible than on the x_{IS} image computed at the iterative solver convergence. Hence, we train a CNN to learn the transformation mapping the early solution x_{RIS} to the corresponding x_{IS}.

Formally, we denote with $\mathcal{T}_j$ the function describing the action of the j-th iteration of the solver:

$$x^{(j+1)} = \mathcal{T}_j(x^{(j)}; b) \quad \forall j \geq 0.$$

Thus, the whole iterative process can be expressed as the concatenation of the following g_K and f_K functions:

$$g_K := \mathcal{T}_{K-1} \circ \cdots \circ \mathcal{T}_0 \tag{7}$$

and

$$f_K := \mathcal{T}_{K^*} \circ \cdots \circ \mathcal{T}_{K+1} \circ \mathcal{T}_K. \tag{8}$$

where K^* is the number iterations necessary to achieve the x_{IS} solution. In fact, it holds:

$$f_K(g_K(x^{(0)}; b)) = x_{IS}. \tag{9}$$

Since we compute $x_{RIS} = g_K(x^{(0)}; b)$ in the RIS step, in the ING phase we approximate f_K by learning the map from x_{RIS} to x_{IS} with a CNN. Ideally, if the network would perfectly learn f_K, its output x_{RISING} should be equal to x_{IS}. It is known that this is not provable in practice; nevertheless, we show numerical evidence through very accurate RISING reconstructions on simulations, in section 5.

4. Experimental design and implementation notes

To numerically verify the feasibility of the RISING workflow, we develop numerical simulations on a data set of images with geometric elements, where measures of merits can be computed, as well as real on a data set of medical images, to properly understand the potential of RISING for clinical applications and settings. We generate the sinograms by projecting the images of the data sets according to 2D fan-beam geometries, and by adding white Gaussian noise to the projections.

This section describes the experimental design, whereas the results are reported and discussed in section 5.

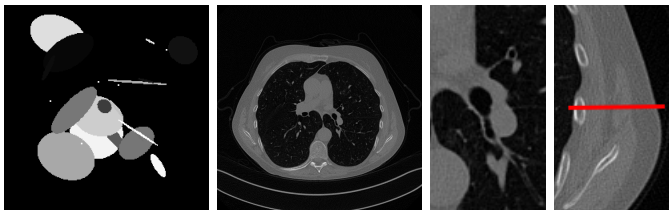

Figure 2: Ground-truth images from the COULE (on the left) and the Low Dose Mayo (center) data sets, with two zoomed crops (on the right) on regions with different anatomical structures.

4.1. Data set of synthetic images

To fully exploit the full-reference image quality assessment metrics and validate our experiments, we created synthetic images and built few-view CT simulations, whose results are analysed in sections 5.1 and 5.2. In particular, the

Constrasted Overlapping Uniform Lines and Ellipses (COULE) data set contains 430 sparse-gradient gray-scale images of size 256×256 with many overlying objects, varying in size and contrast with respect to the background. The left image of Figure 2 shows one image of the data set as an example. The whole data set is downloadable from `www.kaggle.com/loiboresearchgroup/coule-dataset`.

To address sparse-view CT reconstructions, we considered different protocols: the first one is a full angular acquisition with 1-degree spaced projections (we call it $P_{360,360}$ in the following); in the others we reduce the number of acquired projections to 180 ($P_{360,180}$) and 60 ($P_{360,60}$).

4.2. Data set of real medical images

As real patient images, we have downloaded the widely used AAPM Low Dose CT Grand Challenge data set by the Mayo Clinic [36]. It contains 512×512 pixel images of human abdomen, reconstructed from full-dose acquisitions. In Figure 2 we depict one image with two zooms-in highlighting areas with different anatomical structures, such as pulmonary details, sections of ribs and low-contrast inter-costal muscles. On this data set we test the $P_{360,360}$ protocol and a very sparse geometry with only 60 projections in 180 degrees scanning trajectory ($P_{180,60}$). The corresponding results are presented in sections 5.1 and 5.3.

As observable from Figure 2, real full-dose medical images still present little noise and slightly visible streaking artifacts. This makes quite difficult to compute reliable metrics on the reconstructions by referencing them as ground truth.

4.3. Network architecture and training

As CNN architecture we use the state-of-the-art ResUNet architecture, depicted in Figure 3. To approximate the complex image-to-image transformation from x_{RIS} to x_{IS}, we exploit the pooling/unpooling strategies characterizing the popular U-net architecture by [37], where the addition of residual connections between the different levels of image resolution helps maintaining stable the training process. It has already been successfully applied to few-view CT image enhancement in [9, 38]. The activation function is $ReLU$ for each layer, except for the latter where tanh is used.

For the COULE data set, we use $N = 400$ images for training and 30 for testing, whereas for the Mayo Clinic data set we select $N = 3306$ images (corresponding to 10 patients) for the training phase and 357 images (one patient) for testing. Given the data set $\{(x_{RIS,i}, x_{IS,i})\}_{i=1,\dots,N}$, we denote by $F_{\theta^*}(x_{RIS,i})$ the action of the neural network on the input $x_{RIS,i}$ and by $x_{RISING,i} = F_{\theta^*}(x_{RIS,i})$ the network output. We estimate the parameters θ^* by network training, where the loss function is set as $\ell(x_{IS,i}, x_{RISING,i}) = ||x_{RISING,i} - x_{IS,i}||_2^2$, so that:

$$\theta^* = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} \ell(x_{IS,i}, x_{RISING,i}). \tag{10}$$

The training is performed by running Adam for 100 and 50 epochs for the COULE and Mayo datasets, respectively. The batch size is fixed equal to 8 in all

the experiments (this is the largest batch size usable in our Nvidia RTX A4000 GPU). The step size for the optimization algorithm decreases with polynomial decay, going from 10^{-3} to 10^{-5}. To increase the stability over the first iterations, we clip the gradient to 5.

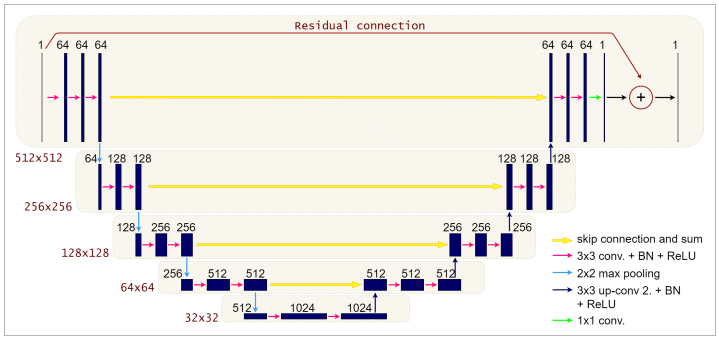

Figure 3: Draft of the ResUNet architecture, implementing the ING task. It is a fully-convolutional neural network with a symmetric encoder-decoder structure and pooling-unpooling operators to enlarge the receptive field.

4.4. Implementation notes

As already presented in section 3.1, we use SGP as iterative solver. In each iteration, we use the class OpTomo of ASTRA toolbox [39, 40] to compute the matrix-vector products Ax and $A^T y$, where A is the matrix in equation (1). In our experiments, we set the smoothing parameter for the TV regularizer to $\beta = 10^{-3}$, the regularization parameter as $\lambda = 10^{-5}$ for the Mayo data set and $\lambda = 4 \cdot 10^{-5}$ for the COULE data set.

To generate the images x_{IS}, we consider the following stopping conditions:

$$\begin{cases} ||\nabla \mathcal{J}(x^{(j)})||_2 < \tau_1 ||\nabla \mathcal{J}(x^{(0)})||_2 \\ ||x^{(j)} - x^{(j-1)}||_2 < \tau_2 ||x^{(j-1)}||_2 \end{cases} \tag{11}$$

where $\nabla \mathcal{J}$ is the gradient of our objective function (5) and $\tau_1 = \tau_2 = 10^{-6}$ in the experiments. All the SGP inner parameters are taken from [34].

4.4.1. Metrics for image quality assessment

To evaluate the efficiency and the accuracy of the reconstructed images, in the following we compute some full-reference image quality assessment metrics. For each considered image x, we consider the relative error (RE) of x with respect to the ground truth x_{GT}:

$$RE(x) = \frac{\|x - x_{GT}\|_2^2}{\|x_{GT}\|_2^2} \tag{12}$$

and the well-known Structural Similarity (SSIM) index, measuring the perceptual difference between two images x and y, as defined in [41]. In our experiments, we compute the SSIM fixing $y = x_{GT}$.

5. Experimental results and discussion

In this section, we report and discuss the more representative numerical experiments performed using the proposed workflow for tomographic image reconstruction from few-view data.

5.1. Robustness of RISING with respect to data perturbation

We first analyse the robustness of the proposed RISING framework with respect to noise on the sinogram. We compare it with a LPP more traditional algorithm (called FBP-LPP in the following), where the same ResUNet architecture used in RISING is trained to map low quality FBP reconstructions to ground truth images [24]. We obtained a noisy projection vector b from the non-noisy sinogram b_{ex} as:

$$b = b_{ex} + 0.01\frac{||b_{ex}||}{||e||}e \tag{13}$$

where e is white noise with standard deviation equal to one. We trained the newtorks on reconstructions computed from the data b and, to verify the methods robustness, we tested them both on sinograms b and b^ϵ corrupted by further noise:

$$b^\epsilon = b + \epsilon\frac{||b||}{||e||}e \tag{14}$$

Figure 4 shows the boxplots of RE and SSIM values drawn from the results on the COULE data set, with the $P_{360,360}$ geometry and $\epsilon = 0.05$. They show that the quality of the images by the FBP-LPP method highly decreases (especially for the SSIM metric) when we consider the noisy data b^ϵ, from the best to the worse values. The results obtained with RISING framework are much more stable: the RE and SSIM values moderately worsen, and the final scores outperform the FBP-LPP. Moreover, the RISING boxplot length in presence of noise is quite smaller, showing that the method performance has not extreme values on the whole test set. Table 1 reports the mean and standard deviation (StdDev in the table) computed on the FBP reconstruction x_{FBP}, the x_{RIS} low quality image and on the two final outputs $x_{FBP-LPP}$ and x_{RISING}. It confirms that RISING is much more robust than FBP-LPP with respect to noise on the data.
The same information on the Mayo test set with geometry $P_{180,60}$ and $\epsilon = 0.04$ is reported in Table 2. We underline that the SSIM values decrease of about 17% for FBP-LPP and only 9.8% for RISING, when applied to high noise data b^ϵ. This provides on medical images the same outcomes obtained on synthetic data.

The previous findings confirm that the widely used FBP-LPP approach gives very good results when applied to images corrupted by the same noise degradating the training samples: in this case it exploits at its best the information provided by the available ground truth images and it outperforms RISING. On the contrary, when the sinograms are affected by unseen noise (as it is common in real applications), the quality of the RISING reconstructions is way better.

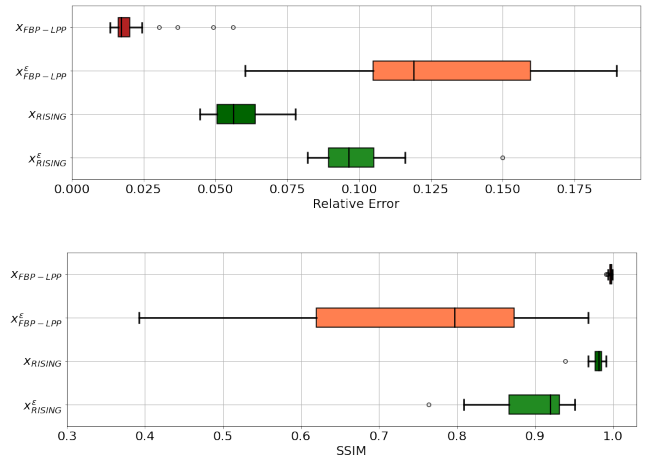

Figure 4: Results on the COULE synthetic test data set, under the $P_{360,360}$ CT protocol. On the x-axis the values of the metrics RE (top) and SSIM (bottom). On the y-axis the framework used from low noise data b (Equation (13)) and high noise data b^{ϵ} (Equation (14))

.

Table 1: Results on the COULE synthetic test data set, under the $P_{360,360}$ CT protocol, reporting mean and standard deviation (StdDev) for different reconstructions obtained with low noise data data b (Equation (13)) and high noise data b^{ϵ} (Equation (14)).

		x_{FBP}		x_{RIS}		$x_{FBP-LPP}$		x_{RISING}	
		Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
RE	b	0.1027	0.0089	0.2501	0.0319	0.0292	0.0097	0.0574	0.0869
	b^{ϵ}	0.2554	0.0132	0.2542	0.0302	0.1294	0.0333	0.0989	0.0128
SSIM	b	0.8926	0.0252	0.8113	0.0686	0.9969	0.0019	0.9801	0.0093
	b^{ϵ}	0.4891	0.0822	0.7916	0.0697	0.7475	0.1537	0.8982	0.0475

5.2. Results on synthetic images

In this paragraph we focus on the synthetic COULE data set.

As first in-depth analysis, we set the geometry $P_{360,360}$ and we compute the x_{RIS} early solution for $K = \{3, 5, 10\}$. Figure 5 shows the x_{RIS} starting images in the top row and the x_{RISING} final reconstructions in the bottom row (relative to the example image of Figure 2). They show that the network is able to almost perfectly restore all three input images. From Table 3 we observe that the x_{RISING} images are very accurate, for $K = \{3, 5, 10\}$, highlighting that the number K of starting iterations seems not to notably influence the final results. Moreover, the table shows that the x_{IS} solution is a reliable target for the network training since it approximates x_{GT} very well.

To understand the behavior of our approach at increasing sparsity in the CT protocol, we also test RISING at different geometric settings such as $P_{360,360}$,

Table 2: Results on the Mayo test data set, under the $P_{180,60}$ CT protocol, reporting mean and standard deviation (StdDev) for different reconstructions obtained with low noise data b (Equation (13)) and high noise data b^ϵ (Equation (14)).

		x_{FBP}		x_{RIS}		$x_{FBP-LPP}$		x_{RISING}	
		Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
RE	b	0.4518	0.0264	0.1803	0.0249	0.1004	0.0164	0.1236	0.0294
	b^ϵ	1.300	0.0609	0.2326	0.0251	0.2026	0.0113	0.2041	0.0229
SSIM	b	0.3672	0.0620	0.8730	0.0283	0.9265	0.0148	0.9145	0.0175
	b^ϵ	0.085	0.0305	0.8016	0.0389	0.7689	0.0681	0.8231	0.0255

Table 3: Mean and standard deviation values of the quality metrics, evaluated on the COULE test set for different RIS reconstructions under the $P_{360,360}$ protocol.

		K=3		K=5		K=10		convergence	
		Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
RE	x_{RIS}	0.6152	0.0366	0.4996	0.0548	0.2700	0.0239	-	-
	x_{RISING}	0.0681	0.0098	0.0784	0.0844	0.0781	0.0122	-	-
	x_{IS}	-	-	-	-	-	-	0.0207	0.0051
SSIM	x_{RIS}	0.2264	0.0262	0.3139	0.0919	0.7844	0.0481	-	-
	x_{RISING}	0.9630	0.0133	0.9267	0.0389	0.9674	0.0140	-	-
	x_{IS}	-	-	-	-	-	-	0.9980	0.0010

Table 4: Mean and standard deviation values of the quality metrics, evaluated on the COULE test set for different geometries.

		$P_{360,360}$		$P_{360,180}$		$P_{360,60}$	
		Mean	StdDev	Mean	StdDev	Mean	StdDev
RE	x_{RIS}	0.2700	0.0239	0.2706	0.0240	0.2834	0.0282
	x_{RISING}	0.0781	0.0122	0.0676	0.0123	0.1113	0.0223
	x_{IS}	0.0207	0.0051	0.0302	0.0077	0.0668	0.0147
SSIM	x_{RIS}	0.7844	0.0481	0.7831	0.0485	0.7472	0.0643
	x_{RISING}	0.9674	0.0140	0.9741	0.0115	0.9493	0.0117
	x_{IS}	0.9980	0.0010	0.9951	0.0030	0.9753	0.0141

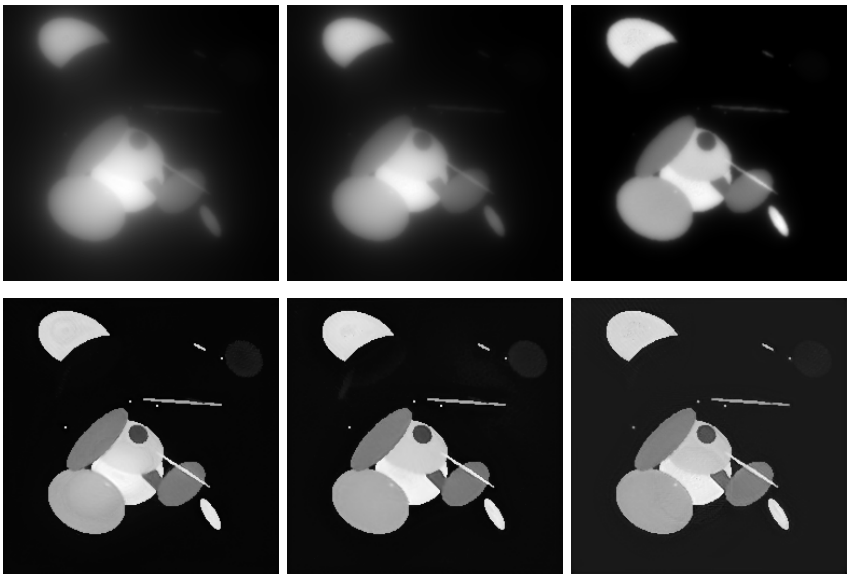

Figure 5: Results on a test image from the COULE synthetic data set, under the $P_{360,360}$ CT protocol. First row, from left to right: x_{RIS} with $K = 3$, $K = 5$ and $K = 10$ respectively; second row, from left to right: the corresponding x_{RISING}.

$P_{360,180}$ and $P_{360,60}$. The number of iterations used to generate x_{RIS} is $K = 10$ and the K^* iterations needed for convergence is in $[150, 300]$ for all the test images. In Table 4 we report the quality indices evaluated on the test set in terms of mean and standard deviation. We first interestingly observe that the values of all the x_{RIS} outputs are very similar, independently from the geometry. Considering the final reconstructions x_{RISING}, we see that halving the number of angles in $P_{360,180}$ does not affect the mean values of the metrics; when the geometry is very sparse in $P_{360,60}$ the errors slightly increase. However, the values of the standard deviations are small and very similar in all the tests, showing that the network has a stable behaviour. As before, the x_{IS} solutions have excellent metrics, justifying the use of the RISING approach even in the hardest case with only 60 angles. This shows that the considered compressed sensing-based model (5) properly describes the reconstruction process independently of the sparsity of the geometry (at least for the considered ones).

5.3. Results on real medical images

Now we present the results of RISING applied to the Mayo data set. The results obtained with two sparse-view CT geometries, namely $P_{360,360}$ and $P_{180,60}$, are shown in Figure 6. The left images are the $x_{RIS} = x^{(15)}$ reconstructions. Even if only a small number of iterations is performed, the main structures of the abdomen are visible; however, the images are still blurry and few streaking artifacts are visible in the bottom image.

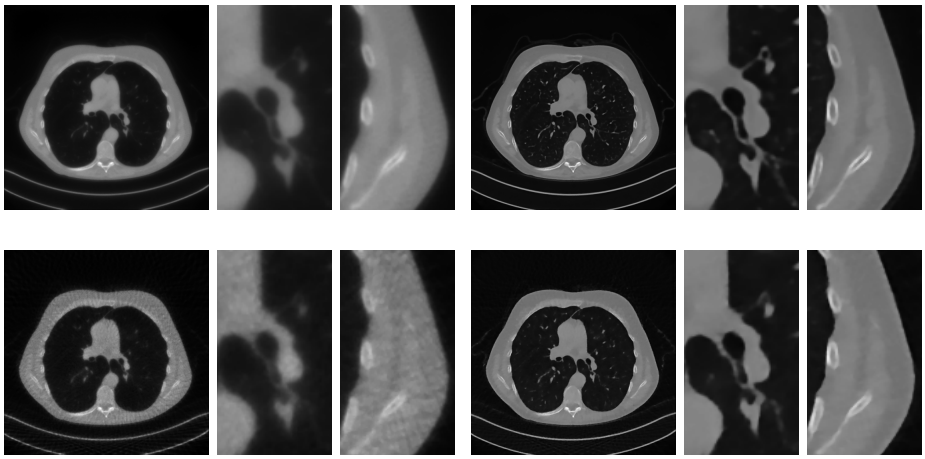

Figure 6: Results on a test image from the Mayo data set, under the $P_{360,360}$ (upper row) and the $P_{180,60}$ (bottom row) CT protocol. Left: x_{RIS}; right: x_{RISING}.

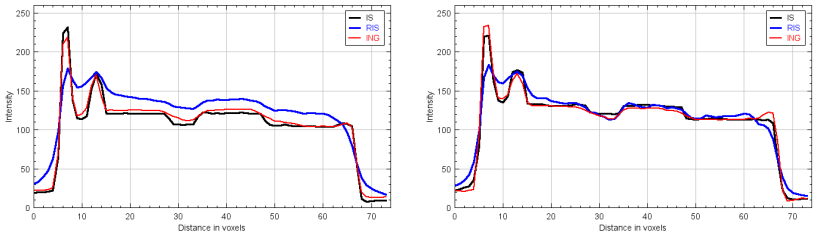

Figure 7: Intensity profiles taken on the horizontal red line depicted in 2, on the reconstructions in Figure 6. The black, blue and red lines are the profile relative to x_{IS}, x_{RIS} and x_{RISING}, respectively. Left $P_{360,360}$; right: $P_{180,60}$.

Our solutions x_{RISING} are reported on the right. They both appear very accurate, the low-contrast regions are correctly preserved and the noise is not visible. In case of the more challenging protocol $P_{180,60}$, the thin details are less neat but still present and discernable. Figure 7 plots the intensity profiles taken over the red line on the crop of Figure 2. In particular, we compare the unseen target profile (black line) from x_{IS} with the network input x_{RIS} curve (blue line) and the output x_{RISING} curve (red line) for both the geometries. In the left image ($P_{360,360}$) we observe that, starting from the input profile, quite far from the black one, we obtain a result faithful to the target, mirroring that the CNN has accurately learnt a map well approximating f_K of (8). Also in the right image ($P_{180,60}$) the line corresponding to the RISING solution almost overlaps the target one. At last we underline that the solutions x_{IS} of the regularized model are very similar in case of $P_{360,360}$ and $P_{180,60}$ geometries.

6. Conclusions

We have proposed a novel framework, called RISING, for the reconstruction of a CT image from few-views. RISING is ground truth free, fast and robust with respect to the noise, hence it is a suitable framework for clinical real usage.

The numerical experiments performed both on synthetic and real medical images show that the RISING reconstructions are visually accurate, even in a very sparse geometry with only 60 views in $[0, 180]$ degrees.

At last, we underline that the RISING scheme is flexible, since it can be set on different model-based iterative reconstruction methods (not limited to the model in (5) and/or to the SGP algorithm). In the future, we intend to study more formal properties of RISING as a stable solver for the CT imaging inverse problem and assess the most suitable network architecture for that task.

Funding

This work has been partially supported by the GNCS - Gruppo Nazionale per il Calcolo Scientifico [*"Apprendimento automatico e tecniche variazionali per la tomografia" INdAM GNCS Project*, grant code CUP_E55F55000270001]

References

[1] J. L. Mueller and S. Siltanen, *Linear and nonlinear inverse problems with practical applications*. SIAM, 2012.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] C. Graff and E. Sidky, "Compressive sensing in medical imaging," *Appl. Opt.*, vol. 54, no. 8, pp. C23–C44, 2015.

[4] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1289–1296, 2018.

[5] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*, pp. 945–948, IEEE, 2013.

[6] G. Chen, X. Hong, Q. Ding, Y. Zhang, H. Chen, S. Fu, Y. Zhao, X. Zhang, H. Ji, G. Wang, *et al.*, "Airnet: Fused analytical and iterative reconstruction with deep neural network regularization for sparse-data ct," *Medical physics*, vol. 47, no. 7, pp. 2916–2930, 2020.

[7] P. Cascarano, E. L. Piccolomini, E. Morotti, and A. Sebastiani, "Plug-and-play gradient-based denoisers applied to ct image enhancement," *Applied Mathematics and Computation*, vol. 422, p. 126967, 2022.

[8] Y. S. Han, J. Yoo, and J. C. Ye, "Deep residual learning for compressed sensing ct reconstruction via persistent homology analysis," *arXiv preprint arXiv:1611.06391*, 2016.

[9] Y. Han and J. C. Ye, "Framing u-net via deep convolutional framelets: Application to sparse-view ct," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1418–1429, 2018.

[10] R. Cavicchioli, J. Hu, E. Loli Piccolomini, E. Morotti, and L. Zanni, "A first-order primal-dual algorithm for convex problems with applications to imaging. GPU acceleration of a model-based iterative method for digital breast tomosynthesis," *Scientific Reports*, vol. 10, no. 1, pp. 120–145, 2020.

[11] E. L. Piccolomini and E. Morotti, "A model-based optimization framework for iterative digital breast tomosynthesis image reconstruction," *Journal of Imaging*, vol. 7, no. 2, p. 36, 2021.

[12] E. Y. Sidky, I. Lorente, J. G. Brankov, and X. Pan, "Do cnns solve the ct inverse problem?," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 6, pp. 1799–1810, 2020.

[13] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.

[14] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. `http://www.deeplearningbook.org`.

[15] J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Problems*, vol. 33, no. 12, p. 124007, 2017.

[16] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1322–1332, 2018.

[17] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "Cnn-based projected gradient descent for consistent ct image reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1440–1453, 2018.

[18] J. Xiang, Y. Dong, and Y. Yang, "Fista-net: Learning a fast iterative shrinkage thresholding network for inverse problems in imaging," *IEEE Transactions on Medical Imaging*, vol. 40, no. 5, pp. 1329–1339, 2021.

[19] H. Zhang, B. Liu, H. Yu, and B. Dong, "Metainv-net: Meta inversion network for sparse view ct image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 40, no. 2, pp. 621–634, 2020.

[20] D. M. Pelt, K. J. Batenburg, and J. A. Sethian, "Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks," *Journal of Imaging*, vol. 4, no. 11, p. 128, 2018.

[21] T. Zhang, H. Gao, Y. Xing, Z. Chen, and L. Zhang, "Dualres-unet: Limited angle artifact reduction for computed tomography," in *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–3, IEEE, 2019.

[22] A.-K. Schnurr, K. Chung, T. Russ, L. R. Schad, and F. G. Zöllner, "Simulation-based deep artifact correction with convolutional neural networks for limited angle artifacts," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 150–161, 2019.

[23] Y. Urase, M. Nishio, Y. Ueno, A. K. Kono, K. Sofue, T. Kanda, T. Maeda, M. Nogami, M. Hori, and T. Murakami, "Simulation study of low-dose sparse-sampling ct with deep learning-based reconstruction: usefulness for evaluation of ovarian cancer metastasis," *Applied Sciences*, vol. 10, no. 13, p. 4446, 2020.

[24] E. Morotti, D. Evangelista, and E. Loli Piccolomini, "A green prospective for learned post-processing in sparse-view tomographic reconstruction," *Journal of Imaging*, vol. 7, no. 8, p. 139, 2021.

[25] T. A. Bubba, G. Kutyniok, M. Lassas, M. Maerz, W. Samek, S. Siltanen, and V. Srinivasan, "Learning the invisible: a hybrid deep learning-shearlet framework for limited angle computed tomography," *Inverse Problems*, vol. 35, no. 6, p. 064002, 2019.

[26] Z. Jiang, Y. Chen, Y. Zhang, Y. Ge, F.-F. Yin, and L. Ren, "Augmentation of cbct reconstructed from under-sampled projections using deep learning," *IEEE transactions on medical imaging*, vol. 38, no. 11, pp. 2705–2715, 2019.

[27] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.

[28] E. L. Piccolomini and E. Morotti, "A fast total variation-based iterative algorithm for digital breast tomosynthesis image reconstruction," *Journal of Algorithms & Computational Technology*, vol. 10, no. 4, pp. 277–289, 2016.

[29] Q. Xu, D. Yang, J. Tan, A. Sawatzky, and M. A. Anastasio, "Accelerated fast iterative shrinkage thresholding algorithms for sparsity-regularized cone-beam ct image reconstruction," *Medical physics*, vol. 43, no. 4, pp. 1849–1872, 2016.

[30] E. Y. Sidky, R. Chartrand, J. M. Boone, and X. Pan, "Constrained TpV minimization for enhanced exploitation of gradient sparsity: Application to ct image reconstruction," *IEEE journal of translational engineering in health and medicine*, vol. 2, pp. 1–18, 2014.

[31] Z. Purisha, J. Rimpeläinen, T. Bubba, and S. Siltanen, "Controlled wavelet domain sparsity for x-ray tomography," *Measurement Science and Technology*, vol. 29, no. 1, p. 014002, 2017.

[32] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259 – 268, 1992.

[33] S. Bonettini, R. Zanella, and L. Zanni, "A scaled gradient projection method for constrained image deblurring," *Inverse problems*, vol. 25, no. 1, p. 015002, 2008.

[34] E. Loli Piccolomini, V. Coli, E. Morotti, and L. Zanni, "Reconstruction of 3D X-ray CT images from reduced sampling by a scaled gradient projection algorithm," *Comp. Opt. Appl.*, vol. 71, pp. 171–191, 2018.

[35] S. Bonettini, F. Porta, and V. Ruggiero, "A variable metric inertial method for convex optimization," *SIAM J. Sci. Comput*, vol. 31, no. 4, pp. A2558–A2584, 2016.

[36] C. McCollough, "Tu-fg-207a-04: Overview of the low dose ct grand challenge," *Medical physics*, vol. 43, no. 6Part35, pp. 3759–3760, 2016.

[37] 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*, pp. 234–241, Springer, 2015.

[38] J. C. Ye, Y. Han, and E. Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *SIAM Journal on Imaging Sciences*, vol. 11, no. 2, pp. 991–1048, 2018.

[39] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The astra toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015.

[40] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. D. Beenhouwer, K. J. Batenburg, and J. Sijbers, "Fast and flexible x-ray tomography using the astra toolbox," *Opt. Express*, vol. 24, pp. 25129–25147, Oct 2016.

[41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.