Plug-and-Play gradient-based denoisers applied to CT image enhancement

Published Version:
Plug-and-Play gradient-based denoisers applied to CT image enhancement / Pasquale Cascarano; Elena Loli Piccolomini; Elena Morotti; Andrea Sebastiani. - In: APPLIED MATHEMATICS AND COMPUTATION. - ISSN 0096-3003. - STAMPA. - 422:(2022), pp. 126967.1-126967.15. [10.1016/j.amc.2022.126967]

Availability:
This version is available at: https://hdl.handle.net/11585/903602 since: 2022-11-18

Published:
DOI: http://doi.org/10.1016/j.amc.2022.126967

Plug-and-Play gradient-based denoisers applied to CT image enhancement

Pasquale Cascarano[a], Elena Loli Piccolomini[b,*], Elena Morotti[c], Andrea Sebastiani[a]

[a]*Department of Mathematics, University of Bologna, Piazza di Porta S. Donato 5, 40126 Bologna, Italy*
[b]*Department of Computer Science and Engineering, University of Bologna, Via Mura Anteo Zamboni 7, 40126 Bologna, Italy*
[c]*Department of Political and Social Sciences, University of Bologna, Strada Maggiore 45, 40125 Bologna, Italy*

Abstract

Blur and noise corrupting Computed Tomography (CT) images can hide or distort small but important details, negatively affecting the consequent diagnosis. In this paper, we present a novel gradient-based Plug-and-Play (PnP) algorithm and we apply it to restore CT images. The plugged denoiser is implemented as a deep Convolutional Neural Network (CNN) trained on the gradient domain (and not on the image one, as in state-of-the-art works) and it induces an external prior onto the restoration model. We further consider a hybrid scheme which combines the gradient-based external denoiser with an internal one, obtained from the Total Variation functional. The proposed frameworks rely on the Half-Quadratic Splitting scheme and we prove a general fixed-point convergence theorem, under weak assumptions on both the denoisers. The experiments confirm the effectiveness of the proposed gradient-based approach in restoring blurred noisy CT images, both in simulated and real medical settings. The obtained performances outperform the achievements of many state-of-the-art methods.

Keywords: Deblur and denoise, Plug-and-Play, gradient-based regularization, external-internal image priors, CNN denoisers, Computed Tomography imaging.

*Corresponding author
Email address: `elena.loli@unibo.it` (Elena Loli Piccolomini)

1. Introduction

In the field of computational imaging, Image Restoration (IR) aims at recovering an unknown clean image from its noisy and/or blurred measurement. In Computed Tomography (CT) the presence of blur and noise reduces diagnostic accuracy, hiding or distorting some small but important objects in the reconstructed image. There are different hardware sources of error which cause blur, such as the finite X-ray focal spot size or the spreading effect in the scintillator in Cone Bean Computed Tomography [1]. Moreover, quantum noise creating random variations in the attenuation coefficients of X-rays, represents the main contribution to the total noise in CT images. Many statistical analysis have shown that the image noise generated by CT scanner can be regarded as normally distributed [2, 3]. Since it is very difficult to avoid these effects by hardware techniques, the software approach is fundamental and several algorithms have been proposed to reduce the blurring and noise artifacts in the CT images. Examples of restoration algorithms for CT images acquired with different geometries can be found in [4, 5, 6, 7] and references therein.

Mathematically, by lexicographically reordering the images as vectors, a generic IR task can be written as the following inverse problem:

$$\text{find } \mathbf{u} \quad \text{such that} \quad \mathbf{v} = \mathbf{A}\mathbf{u} + \mathbf{e}, \tag{1}$$

where $\mathbf{v} \in \mathbb{R}^n$ is the given image, $\mathbf{u} \in \mathbb{R}^n$ is the unknown desired image and $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the forward linear operator defining the IR specific task. The observed image $\mathbf{v}$ is usually affected by noise $\mathbf{e} \in \mathbb{R}^n$, which we assume in this work as Additive White Gaussian Noise (AWGN).

In general, IR problems as (1) are well-known to be ill-posed, meaning that the properties of existence, uniqueness and stability of the desired solution $\mathbf{u}$ are not all guaranteed [8]. Hence, model-based reconstruction methods attempt to find a good estimate $\mathbf{u}^* \in \mathbb{R}^n$ as the solution of a minimization problem whose objective function is the sum of two terms f and g, namely:

$$\mathbf{u}^* \in \arg\min_{\mathbf{u} \in \mathbb{R}^n} \{f(\mathbf{u}) + g(\mathbf{u})\}. \tag{2}$$

The functions f and g are usually referred to as *data fidelity* and *regularization* terms, respectively. The former is a task-related term which models the noise affecting the starting measurement $\mathbf{v}$, whereas the latter induces prior information on the estimate $\mathbf{u}^*$ by reflecting, for example, sparsity patterns, smoothness or geometric assumptions. Often, f is set as an L_p-norm based function measuring the residual between $\mathbf{Au}$ and $\mathbf{v}$, with p strictly related to noise statistics. It is well-known that a squared L_2-norm fidelity fits with the previous assumption of AWGN affecting the measurement $\mathbf{v}$.

The choice of a regularizer is a crucial task in this model-based approach. A widely used strategy is to define g as a handcrafted term based on desired properties of the reconstructed image in a specific domain, such as the gradient or the wavelet domain which have already demonstrated to be effective in medical imaging. In particular, the Total Variation (TV) [9] is largely employed in the IR field for its effectiveness in removing noise and preserving curved contours of the objects [10, 11, 12].

A recent new frontier in the image processing field is represented by the Plug-and-Play (PnP) framework, firstly proposed in [13], where the authors strikingly showed that a closed-form regularizer may not be the best possible choice to properly induce prior information on the desired solution. Technically, the PnP approach derives from the iterative scheme of proximal algorithms, applied to solve regularized optimization problems as (2), whose resulting modular structure allows to deal with the data fidelity f and the regularization term g, separately. Here, in fact, the sub-step involving g reads as a denoising problem, thus it can be replaced by any off-the-shelf denoiser and the computed solution inherits prior information which does not necessarily derive from a closed-form regularization term.

So far, a large number of papers on PnP have been published analyzing different aspects of the scheme, such as the proximal algorithm or the included denoiser. In particular, the considered proximal algorithms are the Alternating Direction Method of Multipliers (ADMM), the Half-Quadratic Splitting (HQS) or the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [13, 14, 15, 16]. In the last few years, PnP methods have been analyzed both in the consensus equilibrium (CE) approach and in the learning to optimize (L2O) framework [17, 18].

Focusing on the choice of the plugged denoiser, several proposals have already been successfully tested and they are usually labelled as internal or external denoisers [19]. Internal denoisers are tailored to define features onto the observed data and they thus induce *internal priors* onto the restored images. As consequence, they struggle to deal with several different image features simultaneously. Examples are the proximal maps of handcrafted regularizers, the BM3D [20] and the Non-Local Mean (NLM) filter [21]. External denoisers are related to an outer set of clean images, so they can fail when dealing with unseen noise variance and image patterns. They induce *external priors* on the IR model. Early studies made use of Gaussian Mixture Models (GMMs) [22] and trained nonlinear reaction diffusion based denoisers [23] as external denoisers. Since nowadays deep learning based priors lead to outstanding performances for denoising images [24, 25], PnP frameworks are also equipped with pre-trained Convolutional Neural Network (CNN) denoisers in works such as [15, 26, 27].

The aforementioned approaches exploit either external or internal denoisers; very recently some generalizations to handle multiple internal and/or external denoisers have been proposed in [28, 29].

Motivation and contributions of the paper

Nowadays X-rays CT systems are designed to acquire images of almost every part of the human body. As a result, tomographic images are quite different from each other and they may contain several objects of various size, shape, and contrast with respect to the background. Moreover, the objective of the imaging task can be to identify one object, would it be small and contrasted as a breast microcalcification, a low-contrast tumoral mass, a larger bone with neat edges or a very thin vessel. In some cases, it is also necessary to subsequently segment the object or an area of interest in the restored image, to help the doctors.

It is well known that priors defined on the gradient domain may enhance medical image reconstructions both in terms of shape recovering and noise removal [30, 12]. Interesting, very few works have so far exploited the PnP scheme to restore CT images [31, 32] and, among the wide literature of PnP, the embedded deep learning based denoisers have always considered only the image space.

This work proposes a PnP framework specifying a gradient-based CNN prior, to solve the different restoration tasks which typically occur in CT medical imaging through CNN networks trained to restore the corrupted image gradients. Moreover, motivated by the apparent complementarity of external and internal denoisers, we also propose a hybrid PnP scheme combining the Total Variation and our CNN-based denoiser. The considered PnP frameworks rely on the Half-Quadratic Splitting algorithm: we derive a fixed point convergence proof upon weak assumptions on the considered denoisers.

We test the methods to restore blurred and noisy synthetic and real CT images. The performances of our proposals are validated through comparisons with other state-of-the-art PnP methods exploiting different denoisers. The numerical results provide very high quality reconstructions and confirm the robustness of the proposed gradient-based frameworks both in restoring different objects of CT images and in removing noise.

Organization of the paper

In this paper, we present in Section 2 the proposed PnP methods together with some implementation choices for the considered denoisers. In Section 3 we report and analyse the numerical results. Finally, in Section 4, we conclude the paper with a brief discussion. In Appendix A we report a fixed-point convergence theorem for the proposed schemes and its proof.

2. Proposed Plug-and-Play methods

We describe here the proposed algorithm for the solution of problem (2) in the wider case of two priors.

Due to the previous assumption of AWGN affecting the measurement $\mathbf{v}$, we fix the fidelity term as $f(\mathbf{u}) := \frac{1}{2}\|\mathbf{A}\mathbf{u} - \mathbf{v}\|_2^2$. As regularizer, we consider a general setting defining g as the sum of two terms g_1 and g_2, weighted by the nonnegative parameters λ and η, respectively. We state the problem assuming that g_1 and g_2 act on the unknown image $\mathbf{u}$ by means of operators $\mathbf{L}_1$ and $\mathbf{L}_2$, respectively. In this case, the minimization problem (2) reads:

$$\mathbf{u}^* \in \arg\min_{\mathbf{u}\in\mathbb{R}^n} \left\{ \frac{1}{2}\|\mathbf{A}\mathbf{u} - \mathbf{v}\|_2^2 + \lambda g_1(\mathbf{L}_1\mathbf{u}) + \eta g_2(\mathbf{L}_2\mathbf{u}) \right\}. \tag{3}$$

where we assume g_1 and g_2 positive and convex real-valued maps:

$$g_1 : \mathbb{R}^{l_1} \to \mathbb{R}^+, \quad g_2 : \mathbb{R}^{l_2} \to \mathbb{R}^+, \tag{4}$$

with l_1 and l_2 positive integers, $\mathbf{L}_1 \in \mathbb{R}^{l_1 \times n}$ and $\mathbf{L}_2 \in \mathbb{R}^{l_2 \times n}$.

We now consider the HQS iterative method described in [33, 34] as numerical solver to compute $\mathbf{u}^*$. By introducing the auxiliary variables $\mathbf{t} \in \mathbb{R}^{l_1}$ and $\mathbf{z} \in \mathbb{R}^{l_2}$ subject to $\mathbf{t} := \mathbf{L}_1\mathbf{u}$ and $\mathbf{z} := \mathbf{L}_2\mathbf{u}$, the following penalized half-quadratic function is taken into account:

$$\mathcal{L}(\mathbf{u},\mathbf{t},\mathbf{z};\rho^{\mathbf{t}},\rho^{\mathbf{z}}) := \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{v}\|_2^2 + \lambda g_1(\mathbf{t}) + \eta g_2(\mathbf{z}) + \\ + \frac{\rho^{\mathbf{t}}}{2}\|\mathbf{L}_1\mathbf{u}-\mathbf{t}\|_2^2 + \frac{\rho^{\mathbf{z}}}{2}\|\mathbf{L}_2\mathbf{u}-\mathbf{z}\|_2^2. \tag{5}$$

At each iteration k, the HQS algorithm performs this alternated minimization scheme with respect to $\mathbf{t}$, $\mathbf{z}$ and the primal variable $\mathbf{u}$:

$$\begin{cases} \mathbf{t}_{k+1} \in \arg\min_{\mathbf{t}\in\mathbb{R}^{l_1}} \lambda g_1(\mathbf{t}) + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathbf{L}_1\mathbf{u}_k - \mathbf{t}\|_2^2 & (6) \\ \mathbf{z}_{k+1} \in \arg\min_{\mathbf{z}\in\mathbb{R}^{l_2}} \eta g_2(\mathbf{z}) + \frac{\rho_k^{\mathbf{z}}}{2}\|\mathbf{L}_2\mathbf{u}_k - \mathbf{z}\|_2^2 & (7) \\ \mathbf{u}_{k+1} = \arg\min_{\mathbf{u}\in\mathbb{R}^{n}} \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathbf{L}_1\mathbf{u} - \mathbf{t}_{k+1}\|_2^2 + \frac{\rho_k^{\mathbf{z}}}{2}\|\mathbf{L}_2\mathbf{u} - \mathbf{z}_{k+1}\|_2^2, & (8) \end{cases}$$

where $(\rho_k^{\mathbf{t}})_{k=1}^{\infty}$ and $(\rho_k^{\mathbf{z}})_{k=1}^{\infty}$ are two non-decreasing sequences of positive penalty parameters. The key feature of HQS is that the prior related substeps (6) and (7) are specified through the proximal maps of g_1 and g_2, respectively, which are mathematically equivalent to regularized denoising problems. The PnP framework exploits both this equivalence and the modular structure of the algorithm by replacing such proximal maps with any off-the-shelf denoiser.

To define the hybrid PnP scheme, we introduce a pre-trained learning-based denoiser $\mathcal{D}_{\sigma}^{\mathrm{ext}}$ and an image-specific denoiser $\mathcal{D}_{\gamma}^{\mathrm{int}}$. These denoisers depend on the positive parameters σ and γ which are related to the noise-level in the images to denoise, so that the greater σ and γ, the stronger the denoising effect is. In particular, in our scheme we choose two sequences $(\sigma_k)_{k=1}^{+\infty}$ and $(\gamma_k)_{k=1}^{+\infty}$ such that, at step k, $\mathcal{D}_{\sigma_k}^{\mathrm{ext}}$ and $\mathcal{D}_{\gamma_k}^{\mathrm{int}}$ replace the sub-steps (6) and (7), respectively. A standard assumption in PnP is that σ_k and γ_k are both related with the penalty parameters $\rho_k^{\mathbf{t}}$ and $\rho_k^{\mathbf{z}}$ through these formulas:

$$\sigma_k := \sqrt{\frac{\alpha}{\rho_k^{\mathbf{t}}}}, \quad \gamma_k := \sqrt{\frac{\beta}{\rho_k^{\mathbf{z}}}}, \tag{9}$$

where α and β are chosen positive scaling factors. A sketch of the resulting hybrid PnP framework is reported in Algorithm 1.

Algorithm 1 Hybrid PnP HQS scheme

Input: α, β and $(\rho_k^{\mathbf{t}})_{k=1}^{\infty}$, $(\rho_k^{\mathbf{z}})_{k=1}^{\infty}$, $\mathbf{A}$, $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{v}$, $\mathbf{u}_1$, K.
for k = 1 … K **do**
  $\mathbf{t}_{k+1} = \mathcal{D}_{\sigma_k}^{\text{ext}}(\mathbf{L}_1\mathbf{u}_k)$
  $\mathbf{z}_{k+1} = \mathcal{D}_{\gamma_k}^{\text{int}}(\mathbf{L}_2\mathbf{u}_k)$
  $\mathbf{u}_{k+1} = \arg\min_{\mathbf{u}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathbf{L}_1\mathbf{u}-\mathbf{t}_{k+1}\|_2^2 + \frac{\rho_k^{\mathbf{z}}}{2}\|\mathbf{L}_2\mathbf{u}-\mathbf{z}_{k+1}\|_2^2$
end for

We remark that under some quite general assumptions on the denoisers and on the sequences $\rho_k^{\mathbf{t}}$ and $\rho_k^{\mathbf{z}}$, the iterates defined in Algorithm 1 converge to a fixed-point (u^*, t^*, z^*). In the Appendix A, an in-depth discussion on the hypothesis and the fixed-point convergence theorem are reported.

As regards the choice of the external denoiser, due to the state-of-the-art performances in denoising tasks reached by deep learning strategies [24, 25], we embed a deep CNN denoiser $\mathcal{D}_{\sigma}^{\text{CNN}}$ as $\mathcal{D}_{\sigma}^{\text{ext}}$. Previous studies have already successfully inspected a CNN-based PnP [15, 27] whose CNN denoisers act directly only on the image-domain. Conversely, our denoiser acts on the image through an operator $\mathbf{L}_1$, which we set equal to the discrete gradient $\mathbf{D} = (\mathbf{D}_h; \mathbf{D}_v)$, where $\mathbf{D}_h, \mathbf{D}_v \in \mathbb{R}^{n\times n}$ are the finite differences discretization of first order derivative operators along the horizontal and vertical axes, respectively. To investigate the effectiveness provided by the proposed learnt gradient-based prior, we consider the case where only the external denoiser is plugged in (thus excluding the internal prior): we label this scheme as GCNN. We will explain in 2.1 how we have implemented the action of the CNN with respect to the choice of the operator $\mathbf{L}_1$. The general scheme of GCNN is in Algorithm 2.

Algorithm 2 GCNN.

Input: α and $(\rho_k^{\mathbf{t}})_{k=1}^{\infty}$, $\mathbf{A}$, $\mathbf{v}$, $\mathbf{u}_1$, K.
for k = 1 … K **do**
  $\mathbf{t}_{k+1} = \mathcal{D}_{\sigma_k}^{\text{CNN}}(\mathbf{D}\mathbf{u}_k)$
  $\mathbf{u}_{k+1} = \arg\min_{\mathbf{u}} \left\{ \frac{1}{2}\|\mathbf{A}\mathbf{u}-\mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathbf{D}\mathbf{u}-\mathbf{t}_{k+1}\|_2^2 \right\}$
end for

We fix as internal denoiser a scheme based on the Total Variation (TV)

[9]. The properties of edge preserving and noise suppressing of the TV in many image processing applications are well-established. The TV function is defined as:

$$\mathrm{TV}(\mathbf{u}) := \sum_{i=1}^{n} \|(\mathbf{Du})_i\|_2 = \sum_{i=1}^{n} \left(\sqrt{(\mathbf{D}_h\mathbf{u})_i^2 + (\mathbf{D}_v\mathbf{u})_i^2} \right), \tag{10}$$

where $(\mathbf{Du})_i := ((\mathbf{D}_h\mathbf{u})_i, (\mathbf{D}_v\mathbf{u})_i) \in \mathbb{R}^2$, for $i = 1 \dots n$ denotes the discrete image gradient computed at pixel i along the horizontal and vertical axes, separately. Hence, the function g_2 in (3) is set as:

$$\begin{aligned} g_2 : \quad & \mathbb{R}^{2\times n} \to \mathbb{R} \\ & \mathbf{x} \to \sum_{i=1}^{n} \|\mathbf{x}_i\|_2 \quad \text{with} \quad \mathbf{x}_i \in \mathbb{R}^2, \end{aligned} \tag{11}$$

assuming $\mathbf{L}_2 = \mathbf{D}$. We remark that, in Algorithm 1, $\mathcal{D}^{\mathrm{int}}_{\gamma_k}$ is the proximal map of g_2 with parameter $\gamma_k^2 = \frac{\eta}{\rho_k^{\mathbf{z}}}$.

The method obtained with the described choices of CNN as external denoiser and TV functional as internal denoiser is reported in Algorithm 3. In the following, we will denote it as GCNN-TV.

Algorithm 3 GCNN-TV.

Input: α, β and $(\rho_k^{\mathbf{t}})_{k=1}^{\infty}$, $(\rho_k^{\mathbf{z}})_{k=1}^{\infty}$, $\mathbf{A}$, $\mathbf{v}$, $\mathbf{u}_1$, K.

for k = $1 \dots K$ **do**

$\mathbf{t}_{k+1} = \mathcal{D}^{\mathrm{CNN}}_{\sigma_k}(\mathbf{Du}_k)$

$\mathbf{z}_{k+1} = \mathrm{prox}_{g_2}(\mathbf{Du}_k)$

$\mathbf{u}_{k+1} = \arg\min_{\mathbf{u}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{Au} - \mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathbf{Du} - \mathbf{t}_{k+1}\|_2^2 + \frac{\rho_k^{\mathbf{z}}}{2}\|\mathbf{Du} - \mathbf{z}_{k+1}\|_2^2$

end for

By the way, we remark that when $\mathbf{L}_1 = \mathbf{I}$, Algorithm 2 is equivalent to the approach proposed in [15] and denoted as ICNN in the following, whereas we label ICNN-TV the algorithm obtained by adding the TV internal prior to ICNN (following the pattern of Algorithm 3).

2.1. Implementation notes

We now refer to particular implementation choices when the proposed algorithms are applied to image deblurring, as considered in our numerical experiments. Here, we refer GCNN-TV and ICNN-TV algorithms. At each iteration k, the minimization problem on the primal variable $\mathbf{u}$ is solved by applying the first order optimality conditions leading to the following linear system:

$$(\mathbf{A}^T\mathbf{A} + \rho_k^{\mathbf{t}}\mathbf{L}_1^T\mathbf{L}_1 + \rho_k^{\mathbf{z}}\mathbf{D}^T\mathbf{D})\mathbf{u}_{k+1} = \mathbf{A}^T\mathbf{v} + \rho_k^{\mathbf{t}}\mathbf{L}_1^T\mathbf{t}_{k+1} + \rho_k^{\mathbf{z}}\mathbf{D}^T\mathbf{z}_{k+1}. \quad (12)$$

This linear system (12) is solvable if the coefficient matrix has full-rank, that is if the following condition holds:

$$\mathbf{Ker}(\mathbf{A}^T\mathbf{A}) \cap \mathbf{Ker}(\mathbf{D}^T\mathbf{D}) \cap \mathbf{Ker}(\mathbf{L}_1^T\mathbf{L}_1) = \{\mathbf{0}\}, \quad (13)$$

where by $\mathbf{Ker}$ we denote the null space of a matrix and $\mathbf{0}$ represents the n-dimensional null vector. The condition (13) is satisfied both for $\mathbf{L}_1 = \mathbf{I}$ and for $\mathbf{L}_1 = \mathbf{D}$. Indeed, $\mathbf{A}$ represents a blurring operator, which is a low-pass filter, whereas the regularization matrix $\mathbf{D}$ is a difference operator, i.e. a high-pass filter. The solution of (12) is given by:

$$\mathbf{u}_{k+1} = (\mathbf{A}^T\mathbf{A} + \rho_k^{\mathbf{t}}\mathbf{L}_1^T\mathbf{L}_1 + \rho_k^{\mathbf{z}}\mathbf{D}^T\mathbf{D})^{-1}(\mathbf{A}^T\mathbf{v} + \rho_k^{\mathbf{t}}\mathbf{L}_1^T\mathbf{t}_{k+1} + \rho_k^{\mathbf{z}}\mathbf{D}^T\mathbf{z}_{k+1}). \quad (14)$$

The direct computation of the analytical solution (14) requires the inversion of a high dimensional matrix. By assuming periodic boundary conditions $\mathbf{A}^T\mathbf{A}$, $\mathbf{D}^T\mathbf{D}$ and $\mathbf{L}_1^T\mathbf{L}_1$ are Block Circulant with Circulant Blocks (BCCB) matrices which can be diagonalized by the two dimensional discrete Fourier transform [35]. Hence, the solution of (12) can be efficiently computed using the two dimensional Fast Fourier Transform (FFT) as:

$$\mathbf{u}_{k+1} = \mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}(\mathbf{A})}\mathcal{F}(\mathbf{v}) + \rho_k^{\mathbf{t}}\overline{\mathcal{F}(\mathbf{L}_1)}\mathcal{F}(\mathbf{t}_{k+1}) + \rho_k^{\mathbf{z}}\overline{\mathcal{F}(\mathbf{D})}\mathcal{F}(\mathbf{z}_{k+1})}{\overline{\mathcal{F}(\mathbf{A})}\mathcal{F}(\mathbf{A}) + \rho_k^{\mathbf{t}}\overline{\mathcal{F}(\mathbf{L}_1)}\mathcal{F}(\mathbf{L}_1) + \rho_k^{\mathbf{z}}\overline{\mathcal{F}(\mathbf{D})}\mathcal{F}(\mathbf{D})}\right) \quad (15)$$

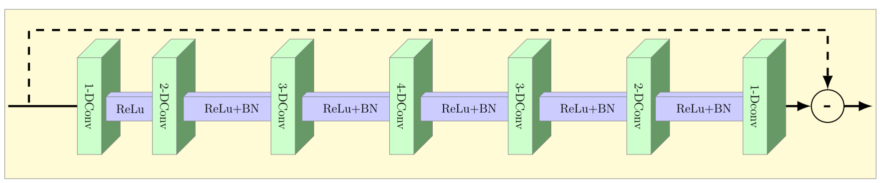
where, for the sake of simplicity, $\mathcal{F}(\cdot)$ and $\overline{\mathcal{F}(\cdot)}$ denote the FFT and its conjugate, whereas $\mathcal{F}^{-1}(\cdot)$ is the inverse FFT. Similarly, $\mathbf{u}_{k+1}$ in Algorithm 2 can be computed by setting $\rho_k^{\mathbf{z}} = 0$ in (15).

Concerning the update of $\mathbf{z}_k$ in Algorithm 3, we observe that it reduces to the solution of n bi-dimensional optimization problems which can be computed in a closed form by using the proximal map of the L_2-norm.

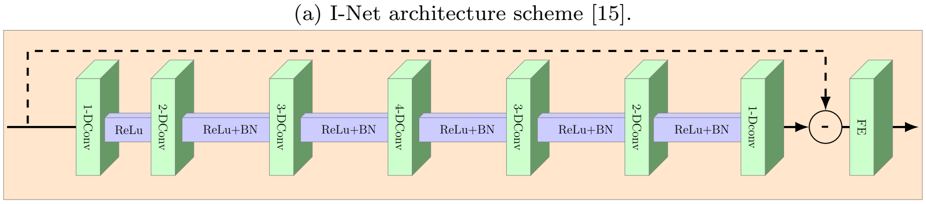
To implement the CNN based external denoiser $\mathcal{D}_\sigma^{\text{CNN}}$ we adopt the widely used DnCNN architecture proposed in [15]. We refer to this architecture, which is shown in Figure 1a, as I-Net. It is constituted by seven dilated

convolutional layers [36] activated by ReLu functions.
For the CNN training, we consider the *Train400* image dataset [23]. It contains 400 gray-scale natural images of size 180×180 obtained by cropping larger images from the Berkeley Segmentation dataset [37]. We make use in our implementation of the 25 denoisers downloaded from `https://github.com/cszn/IRCNN`, each one trained on a single noise level in the range [2, 50] with step 2. As represented in Figure 2a, the I-Net is trained to remove noise from the noisy input images.

Our proposal considers the case $\mathbf{L}_1 = \mathbf{D}$. In this case, we add the linear Feature Extractor (FE) computing the discrete image gradient at the end of the I-Net architecture, thus obtaining the G-Net network depicted in Figure 1b. Therefore, in order to compute the iterate $\mathbf{t}_{k+1}$ as in (6), the G-Net is trained to give as output the noisy-free gradient image taking as input the noisy images as I-Net (Figure 2b). We use the ADAM optimizer with the Tensorflow default parameters and we set the epochs number to 150. The correspondence between the iteration k of the algorithms and one of the 25 available networks is performed as in [15].



(a) I-Net architecture scheme [15].



(b) G-Net architecture scheme.

Figure 1: I-Net and G-Net architecture schemes. BN represents the batch normalization and m-DConv denotes m-dilated convolution.

3. Results and discussion

In this section, we describe the results obtained by testing the proposed schemes on the task of image denoising and deblurring. The Python codes of our proposals are available at `https://github.com/sedaboni/PnP-TV`. We

(a) CNN image denoiser on the image.   (b) CNN denoiser for gradient restoration.

Figure 2: Trained schemes for denoising.

validate our methods both on a synthetic image, characterized by elements of interest for CT medical purposes, and on real CT images. All the ground truth images have values in the range $[0, 255]$.

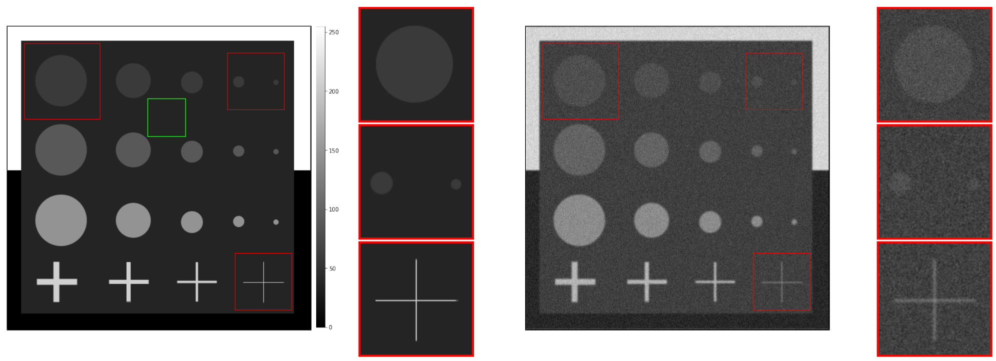
Our methods are compared with the baseline TV regularization implemented in the standard ADMM algorithm, which uses the discrepancy principle [38] for the estimation of the regularization parameter, the approach proposed in [15] which is referred to as ICNN in the following, the standard PnP with BM3D and NLM chosen as denoisers and a very recent method [29] which combines a truncated L_1-norm computed on the wavelet operator applied to the signal and BM3D (BM3D-WL1), therefore two internal denoisers. To complete our comparison, we also consider the ICNN-TV algorithm.

For a quality assessment of the results, we create artificially blurred and noisy images from a ground truth (GT) image and we compute the Structural Similarity Index Measure (SSIM) and the Peak Signal-to-Noise-Ratio (PSNR) [39] between the restored image and the ground-truth. Moreover, to quantify noise removal, we compute the standard deviation on uniform regions of interest of the restored images.

For all the proposed algorithms the input parameters α and β are heuristically chosen to compute a solution satisfying the discrepancy principle. The algorithms perform at most 30 iterations. The first iterate $\mathbf{u}_1$ is initialized as a vector of zeros. Concerning the choice of $(\rho_k^{\mathbf{t}})_{k=1}^{\infty}$ and $(\rho_k^{\mathbf{z}})_{k=1}^{\infty}$, we have set $\rho_k^{\mathbf{t}} = \rho_k^{\mathbf{z}} = k(1+\epsilon)^k$, with $\epsilon > 0$, satisfying the conditions required in the fixed-point convergence theorem stated in Appendix A. All the hyperparameters of the competitors have been fixed in order to provide a solution which satisfies the discrepancy principle.

3.1. Results on a synthetic test problem

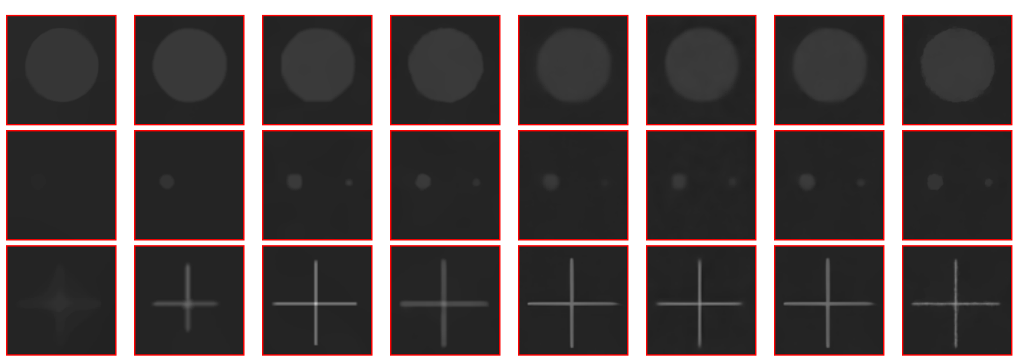
We start our experiments by considering the numerical simulation acting on the gray-scale 512×512 synthetic image reported in Figure 3a. The image is designed to test the algorithms performance in the case of low and



(a) Ground truth (b) Corrupted

Figure 3: Ground truth gray-scale test image and a simulated degraded acquisition. In (a) the green square highlights the uniform patch used to evaluate ROI-std. In (a) and (b) three close-ups (red boxes) are depicted alongside the images.

high contrast objects, with curved and straight borders: the ground truth image contains many circles of different diameter but uniform intensity; each row has homogeneous circles, with enhancing contrast, from top to bottom, with respect to the uniform background. The fourth row contains crosses of different thickness and high contrast. To build our test problems, we blur the ground truth image using a Gaussian 15×15 kernel with zero mean and standard deviation 1.2, then we introduce AWGN with standard deviation std in $\{10, 15, 20\}$. In Figure 3b we show the corrupted image obtained with $std = 15$. In Figure 3a and 3b, we also depict three close-ups on the regions bounded by red squares.



(a) TV (b) NLM (c) BM3D (d) BM3D-WL1 (e) ICNN (f) GCNN (g) ICNN-TV (h) GCNN-TV

Figure 4: Three close-ups for each reconstruction by different methods obtained for the synthetic image.

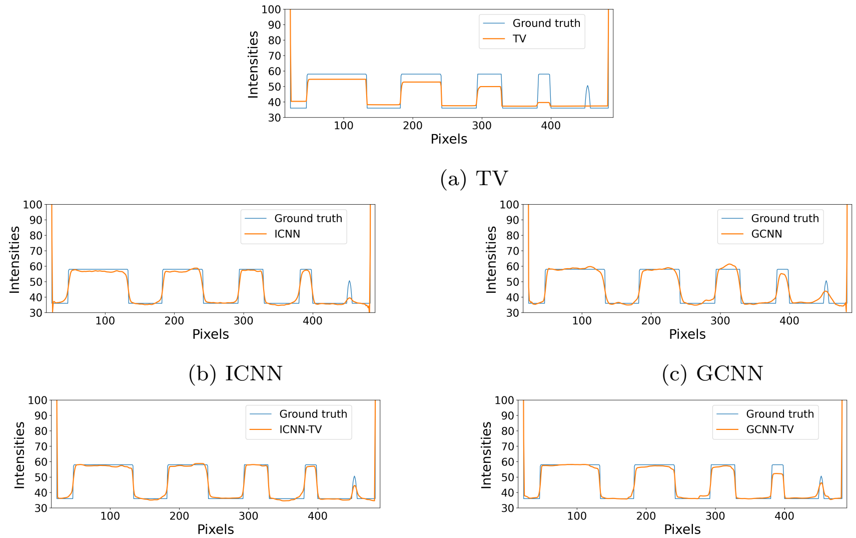
	AWGN of $std = 10$		AWGN of $std = 15$		AWGN of $std = 20$	
	PSNR	ROI-std	ROI-PSNR	ROI-std	PSNR	ROI-std
TV	30.8085	**0.0271**	28.6664	**0.0507**	27.4028	**0.0772**
NLM	32.6266	0.0896	31.3772	0.1122	30.1042	**0.1382**
BM3D	32.1221	0.3657	31.3283	0.5785	30.4806	0.7281
BM3D-WL1	31.7616	0.3974	30.7724	0.5802	30.2951	0.7779
ICNN	34.1838	0.4398	33.0519	0.5555	32.9788	0.7851
GCNN	**34.7078**	0.4749	**33.9640**	0.6568	**33.2446**	0.8189
ICNN-TV	32.3531	0.4081	31.3775	0.4798	30.4499	0.5553
GCNN-TV	33.2648	0.1706	31.7743	0.2512	30.6453	0.3129

Table 1: Measures computed on restored images varying the standard deviation of the AWGN. The two best PSNR and ROI-std (standard deviation computed inside the green square in Figure 3a) values for each AWGN are highlighted in blue and green, respectively. The first best is highlighted in bold.

In Figure 4, for each method we report the three restored zooms in the same range of gray levels. For what concerns the low-contrast circles, reported in the first two rows, it is evident that the hybrid approaches (such as BM3D-WL1, ICNN-TV and GCNN-TV) outperform the other algorithms which exploit only one prior (TV, NLM, BM3D,ICNN, GCNN). Indeed, TV and NLM struggle to retrieve the small circles, whereas BM3D deforms the shape of the objects (Figure 4c). We highlight that the smallest circle is visible in the ICNN reconstruction (Figures 4e and it is further enhanced in the GCNN restoration 4f). Focusing on the restoration of an object, the one-pixel thick cross, with a different shape and contrast, we observe that BM3D, ICNN and GCNN achieve the highest enhancement (see the last row of Figure 4). However we remark that, even in this case, TV and NLM tend to suppress very thin details.

In Figure 5 we plot the pixel intensities of a horizontal image row passing through all the lowest-contrasted circles, to better inspect the effects of adding the TV internal prior to the ICNN and GCNN schemes on the most challenging objects. The plot in Figure 5a reflects the typical loss-of-contrast drawback of the TV prior, oversmoothing the two smallest circles. Adding the TV prior to ICNN and GCNN algorithms removes the residual noise, especially visible in the largest circle, while enhancing the edges.

To test the robustness of the proposed models with respect to the noise, we analyze the results, reported in Table 1, obtained by the considered methods when different variances of the AWGN are considered. We observe that, in terms of PSNR, the GCNN method gets the best values in all the cases, thus confirming the effectiveness of the proposed CNN denoiser defined on



(a) TV

(b) ICNN

(c) GCNN

(d) ICNN-TV

(e) GCNN-TV

Figure 5: Intensity line profiles on the 90th row cutting the lowest contrasted circles. The blue and orange lines represent the ground truth and the restored image profiles for different methods, respectively.

the image gradient domain. When we introduce the contribution of the TV-based internal prior, the PSNR values decrease, even if the global denoising effect due to TV is visually evident, as previously underlined. To confirm this, we report in Table 1 the standard deviation (ROI-std) computed on the constant region marked by the green bounding square in Figure 3a. The TV and NLM methods always have the lowest values, whereas the proposed hybrid approaches ICNN-TV and GCNN-TV are more effective in case of high noise.

3.2. Results on real CT medical images

We now consider two X-ray Computed Tomography images to compare the effectiveness of the proposed schemes. In order to illustrate the advantages of our proposals, according to their features highlighted in the synthetic case, we examine a head and chest CT images containing small and low-contrasted details.

3.2.1. CT head image for epidural hemorrhage detection

The considered head tomographic image is downloaded from an open source dataset[1]. It shows an intracranial hemorrhage, which requires a rapid and intensive medical treatment based on the accurate localization of the blood in the CT image obtained by segmentation algorithms (represented as the red region in Figure 6a). If the image is severely corrupted, the segmentation procedure may fail. As an example, after blurring the ground truth image with a Gaussian kernel of size 15×15 and standard deviation 0.5 and adding AWGN with standard deviation 25, we compute the segmentation mask by an online open source software [2]. The segmented region is shown in red in Figure 6b. To highlight the importance of deblurring and denoising the image before segmenting it, we show the red mask computed on one restored image in Figure 6c.

In Figure 7 we report three close-ups for each method. The first one highlights the central part of the head CT image containing blood vessels, whereas the second zoom shows a portion of the cerebral cortex with sulci. The third zoom of the figure focuses onto the epidural hemorrhage (pointed by the magenta arrow). In Table 2 we report the PSNR computed between the restored image and the ground truth, and the Jaccard similarity coefficient (Jac) between the masks computed on the ground truth and the restored images. By a visual comparison, we observe that TV, NLM, BM3D-WL1 output images look too smooth and blocky whereas the BM3D deforms the anatomical contours. We highlight that the GCNN method accurately restores the vessels and sulci borders and it gets the highest PSNR value, reflecting the effectiveness of the gradient-based regularization. As regard the Jaccard values, the best ones are achieved by the hybrid frameworks (i.e. ICNN-TV and GCNN-TV), where the smoothing effect of the TV-based denoiser improves the border detectability, making the restored images suitable for segmentation tasks.

3.2.2. Restoration of low-dose CT real chest image

We now consider a Computed Tomography open source dataset[3] of real chest images. In Figure 8a we focus on one image (ID: 0005) of the dataset. We point out that it contains many different objects, varying in size, dimen-

[1] `https://www.kaggle.com/vbookshelf/computed-tomography-ct-images`
[2] `http://brain.test.woza.work/`
[3] `https://www.kaggle.com/kmader/siim-medical-images`

(a) GT + mask (b) Corrupted + mask (c) Restored + mask

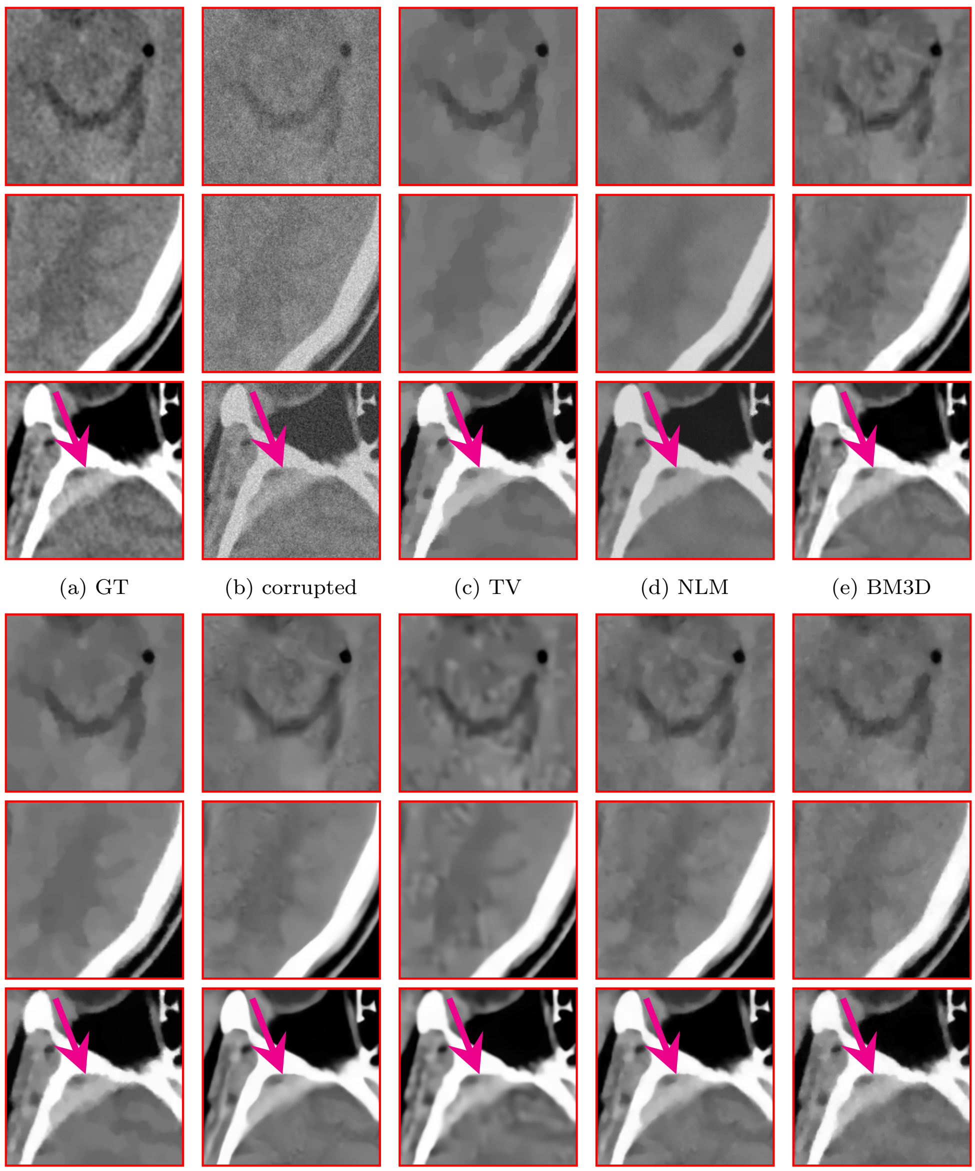
Figure 6: Head tomographic image with epidural hemorrhage. Computed masks are coloured red.

	TV	NLM	BM3D	BM3D-WL1	ICNN	GCNN	ICNN-TV	GCNN-TV
PSNR	30.6781	28.4723	31.7320	31.9917	33.0800	**33.5881**	32.2843	32.0872
Jac	0.9471	0.8827	0.9313	0.9500	0.9387	0.9398	**0.9557**	0.9504

Table 2: PSNR and Jaccard computed on restored image. The two best PSNR and Jaccard values are highlighted in green and blue, respectively. The first best is highlighted in bold.

sion and gray intensity. To simulate a low-dose CT reconstructed image, which is characterized by high noise, after blurring the image by using a Gaussian kernel of dimension 15×15 with standard deviation 0.5, we add AWGN with high standard deviation equals to 25. In Figure 8b we show the very noisy corrupted image where small and low-contrasted details are not well detectable.

In Figure 9 we report three close-ups of the restorations showing different details of the image. In the first close-up we observe that in some cases the borders of the ascending aorta and superior vena cava sections pointed by the arrow are not well distinguishable as in the ground truth image. In particular, we notice that the GCNN method produces the best image. The second crop contains thin vessels immersed in the dark pulmonary background. The images obtained with TV, NML and BM3D-WL1 algorithms are too smooth and some details are hardly visible. In the BM3D and ICNN-based output images the circular sections of the vessels are distorted into triangular shapes, whereas the images obtained with gradient-based CNN restore very well the path of the main vessels, without oversmoothing. In the third row, the close-ups show that only GCNN and GCNN-TV well recover the circular shape of the vertebral canal and GCNN outperforms the competitors in identifying the transverse process edges (Figures 9h and 9j).

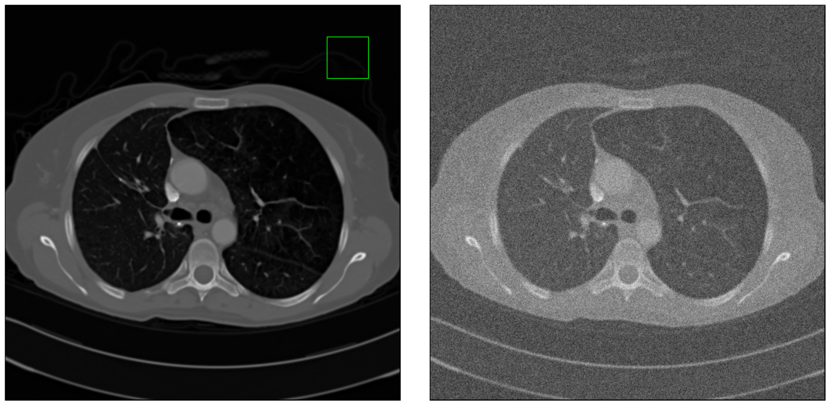


(a) GT (b) corrupted (c) TV (d) NLM (e) BM3D

(f) BM3D-WL1 (g) ICNN (h) GCNN (i) ICNN-TV (j) GCNN-TV

Figure 7: Three close-ups for each reconstruction by different methods obtained for the head CT image. The magenta arrows highlight the epidural hemorrhages.

To deeper analyse the improvement given by the proposed gradient-based CNN over the image-based one, we plot in Figure 10 the profiles relative



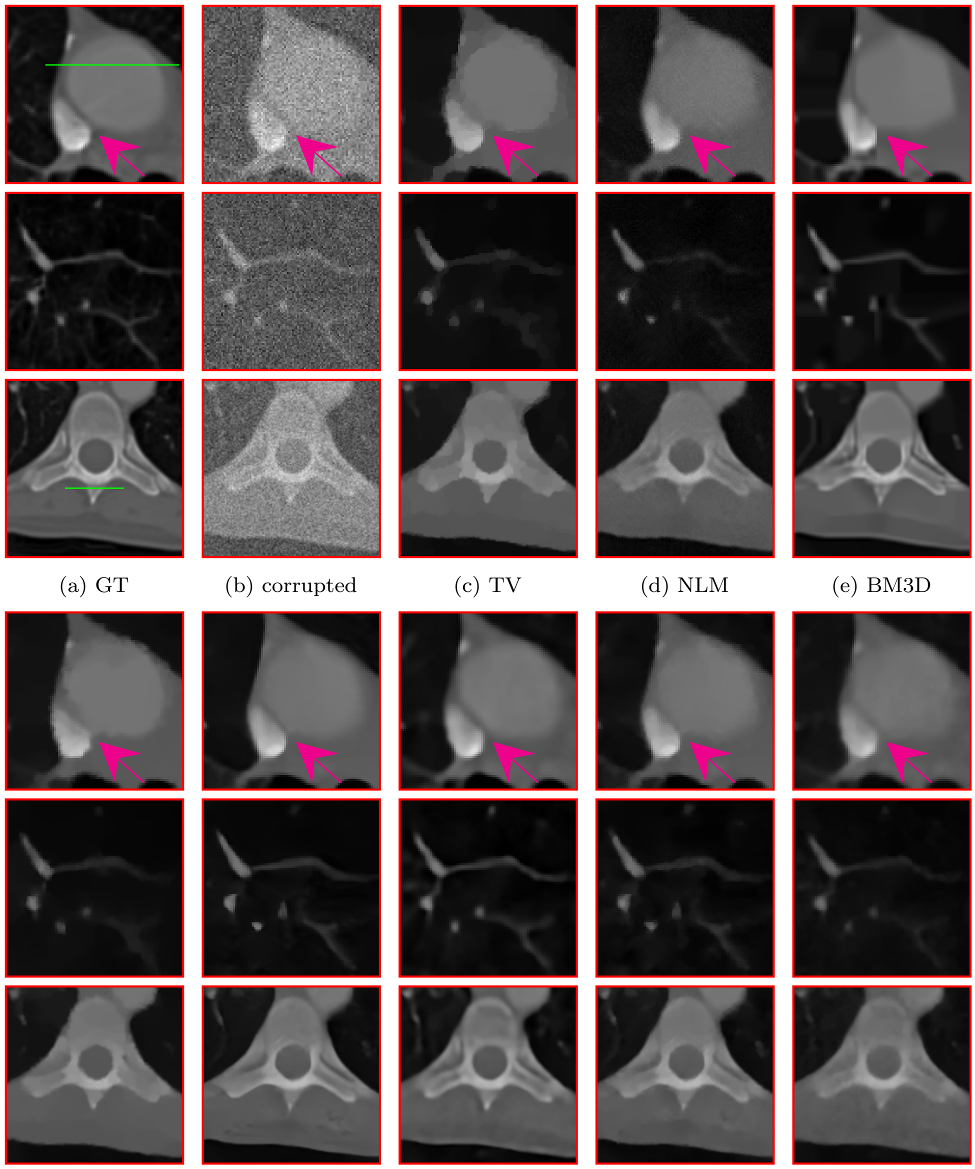
(a) Ground truth (b) Corrupted

Figure 8: Low-dose CT chest image (`ID: 0005`). In (a) the green square highlights the uniform patch used to evaluate ROI-std.

to the green segments depicted in Figure 8a over the first and third crops. The first plot (Figure 10a) refers to a large homogeneous object and it is evident that the GCNN red line better fits the blue line corresponding to the ground truth and that the orange ICNN profile oversmoothes. The profile over the spinous process (Figure 10b) highlights that GCNN better restores thin objects. We can conclude for the restoration of this image that the use of a gradient-based CNN denoiser has advantages such as a better enhancing of the objects contours and the preservation of small details, over the use of an image-based CNN denoiser.

Finally, to measure the reconstruction quality and the residual noise, we compute the PSNR and SSIM measures on the whole image and the standard deviation on a flat region indicated by the green square in Figure 8a. From the Table 3, we observe that the GCNN method attains both the best PSNR and SSIM. The BM3D algorithm achieves the second best PSNR but it often deformates the curve boundary contours of the objects (as in Figure 9e). Regarding the ROI-std measure, as expected, the TV method gets the lowest standard deviation on the region of interest. Moreover, we observe that the addition of TV as internal prior in the CNN-based methods considerably lowers the standard deviation values, as confirmed by ICNN-TV and GCNN-TV columns.

At last, Figure 11 generalises the results of Table 2. We have in fact executed the GCNN and ICNN algorithms on 100 images from the whole

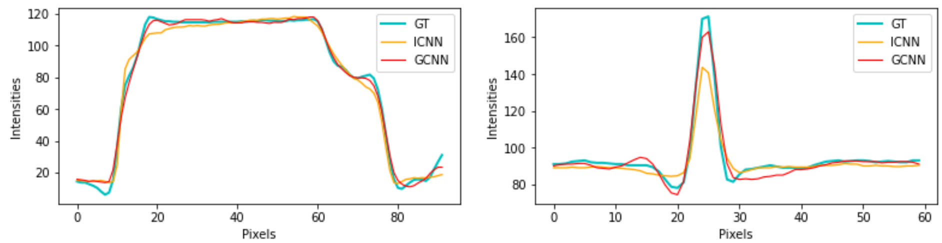


(a) GT (b) corrupted (c) TV (d) NLM (e) BM3D

(f) BM3D-WL1 (g) ICNN (h) GCNN (i) ICNN-TV (j) GCNN-TV

Figure 9: Three close-ups for each reconstruction by different methods obtained for the chest low-dose CT image. The magenta arrows highlight a region of interest.

chest dataset and computed the boxplots relative to the PSNR (Figure 11a) and the SSIM (Figure 11b) metrics. These statistics validate the results



(a) (b)

Figure 10: Intensity line profiles on the horizontal lines depicted in Figure 9a, over the aorta (left) and on the spinous process of the vertebra (right). The blue, orange and red lines represent the ground truth, the ICNN and the GCNN restored image profiles, respectively.

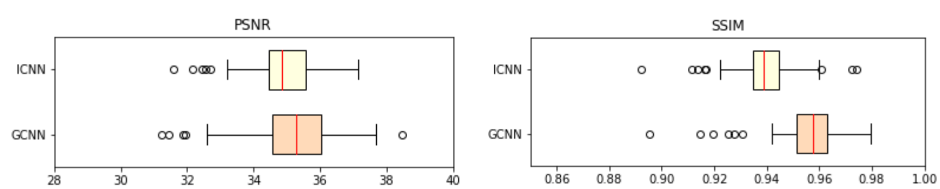
	TV	NLM	BM3D	BM3D-WL1	ICNN	GCNN	ICNN-TV	GCNN-TV
PSNR	32.1727	30.9899	34.7675	32.9104	34.1673	**35.0309**	34.0946	33.5789
SSIM	0.9297	0.9129	0.9499	0.9358	0.9474	**0.9546**	0.9466	0.9443
ROI-std	**0.1746**	0.3017	0.6569	0.5816	1.1136	1.2366	0.2844	0.3460

Table 3: Standard deviation computed on the region of interest inside the green square in Figure 8a, for the Low-Dose CT chest images.

discussed on one single image and confirm that GCNN outperforms ICNN.

4. Conclusions

In this paper we have proposed a new PnP framework using learnt gradient-based priors for CT medical image restoration. We considered a Half-Quadratic Splitting minimization algorithms where the denoising step is executed by a CNN acting on the image gradients (GCNN method). We also considered



(a) (b)

Figure 11: Boxplots of the PSNR values (a) and SSIM values (b) computed on 100 chest images by ICNN algorithm (yellow ones) and GCNN algorithm (orange ones).

a hybrid regularization where we added a Total Variation functional in the GCNN scheme (GCNN-TV).

The numerical experiments on synthetic and real CT medical images show that the proposed GCNN, well recovers the curve contours of flat and low-contrast objects, as well as thin vessels. The obtained image enhancements confirm that gradient-based priors are effective for the restoration of medical CT images, since the competitors get lower quality indices. Indeed, the GCNN-TV further smoothes homogeneous area such as backgrounds and small low-contrast objects on very noisy images and its restoration appears suitable for segmentation.

Acknowledgments

This research was funded by the Indam GNCS grant 2020 *Ottimizzazione per l'apprendimento automatico e apprendimento automatico per l'ottimizzazione.*

A. Appendix

To analyze the convergence properties of Algorithm 1, we start observing that if the denoisers $\mathcal{D}_\sigma^{\text{ext}}$ and $\mathcal{D}_\gamma^{\text{int}}$ are the proximal maps of two convex functions g_1 and g_2, respectively, then the convergence to a global minimum of the objective function in (3) is guaranteed [33, 34]. However, in [14] the authors observe that a denoiser is a proximal map when it is nonexpansive with symmetric gradient, thus limiting the set of suitable denoisers. In the effort of allowing less strict conditions on the involved denoisers, we show in this section that the proposed Algorithm 1 satisfies a fixed-point convergence theorem provided only their boundedness.

Definition A.1 (Bounded Denoiser [40])**.** A *bounded denoiser* with parameter ϵ is a function $\mathcal{D}_\epsilon : \mathbb{R}^l \to \mathbb{R}^l$ such that for any $\mathbf{t} \in \mathbb{R}^l$ the following inequality holds:

$$\|\mathcal{D}_\epsilon(\mathbf{t}) - \mathbf{t}\|_2^2 \le \epsilon^2 C_{\mathcal{D}} \tag{A.1}$$

for a constant $C_{\mathcal{D}}$ independent of ϵ.

The previous definition entails that given the sequence $(\epsilon_k)_{k=1}^{+\infty}$, $\mathcal{D}_{\epsilon_k}$ converges to the identity function of $\mathbb{R}^l$ as $\epsilon_k \to 0$.

In order to state and prove the following fixed-point theorem, we make some assumptions.

Given $(\rho_k^{\mathbf{t}})_{k=1}^{\infty}$ and $(\rho_k^{\mathbf{z}})_{k=1}^{\infty}$ non-decreasing positive sequences, $\mathbf{L}_1 \in \mathbb{R}^{l_1 \times n}$, $\mathbf{L}_2 \in \mathbb{R}^{l_2 \times n}$ as input for Algorithm 1, then we assume:

1. $\mathcal{D}_{\sigma_k}^{\text{ext}}$ and $\mathcal{D}_{\gamma_k}^{\text{int}}$ are bounded denoisers.
2. $\mathbf{L}_1$ and $\mathbf{L}_2$ are full-rank matrices.
3. $\sum_{k=1}^{+\infty} \sqrt{\dfrac{k}{\rho_k^{\mathbf{z}}}} < +\infty, \quad \sum_{k=1}^{+\infty} \sqrt{\dfrac{k}{\rho_k^{\mathbf{t}}}} < +\infty$ and $\dfrac{\rho_k^{\mathbf{z}}}{\rho_k^{\mathbf{t}}} \to c$ where $c \in \mathbb{R}^+$.

Theorem A.1 (Fixed-point convergence theorem for the hybrid PnP algorithm)**.** *Given the assumptions 1-3, there exist* $\mathbf{t}^* \in \mathbb{R}^{l_1}, \mathbf{z}^* \in \mathbb{R}^{l_2}$ *and* $\mathbf{u}^* \in \mathbb{R}^n$ *such that, for* $k \to \infty$*, the following relations hold:*

$$\mathbf{t}_k \to \mathbf{t}^*, \quad \mathbf{L}_1\mathbf{u}_k \to \mathbf{t}^*, \quad \mathbf{z}_k \to \mathbf{z}^*, \quad \mathbf{L}_2\mathbf{u}_k \to \mathbf{z}^*, \quad \mathbf{u}_k \to \mathbf{u}^*,$$

where $\mathbf{t}_k, \mathbf{z}_k, \mathbf{u}_k$ *are computed as in Algorithm 1 at step* k*.*

Proof. By observing that $\mathbf{u}_{k+1}$ is the optimal solution of the minimization problem (8), and by using the relations in (9) and the assumption 1, we get the following chain of inequalities:

$$\begin{aligned}
&\frac{1}{2}\|\mathbf{A}\mathbf{u}_{k+1} - \mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathbf{t}_{k+1} - \mathbf{L}_1\mathbf{u}_{k+1}\|_2^2 + \frac{\rho_k^{\mathbf{z}}}{2}\|\mathbf{z}_{k+1} - \mathbf{L}_2\mathbf{u}_{k+1}\|_2^2 \leq \\
&\leq \frac{1}{2}\|\mathbf{A}\mathbf{u}_k - \mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathbf{t}_{k+1} - \mathbf{L}_1\mathbf{u}_k\|_2^2 + \frac{\rho_k^{\mathbf{z}}}{2}\|\mathbf{z}_{k+1} - \mathbf{L}_2\mathbf{u}_k\|_2^2 = \\
&= \frac{1}{2}\|\mathbf{A}\mathbf{u}_k - \mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\|\mathcal{D}_{\sigma_k}^{\text{ext}}(\mathbf{L}_1\mathbf{u}_k) - \mathbf{L}_1\mathbf{u}_k\|_2^2 + \frac{\rho_k^{\mathbf{z}}}{2}\|\mathcal{D}_{\gamma_k}^{\text{int}}(\mathbf{L}_2\mathbf{u}_k) - \mathbf{L}_2\mathbf{u}_k\|_2^2 \leq \\
&\leq \frac{1}{2}\|\mathbf{A}\mathbf{u}_k - \mathbf{v}\|_2^2 + \frac{\rho_k^{\mathbf{t}}}{2}\sigma_k^2 C_{\mathcal{D}^{\text{ext}}} + \frac{\rho_k^{\mathbf{z}}}{2}\gamma_k^2 C_{\mathcal{D}^{\text{int}}} = \\
&= \frac{1}{2}\|\mathbf{A}\mathbf{u}_k - \mathbf{v}\|_2^2 + \frac{\alpha}{2}C_{\mathcal{D}^{\text{ext}}} + \frac{\beta}{2}C_{\mathcal{D}^{\text{int}}} \leq \\
&= \frac{1}{2}\|\mathbf{A}\mathbf{u}_k - \mathbf{v}\|_2^2 + \tilde{C},
\end{aligned} \tag{A.2}$$

with $\tilde{C} := \dfrac{\alpha}{2}C_{\mathcal{D}^{\text{ext}}} + \dfrac{\beta}{2}C_{\mathcal{D}^{\text{int}}}$.

Since all the considered terms in (A.2) are positive, the following inequalities hold:

$$\frac{1}{2}\|\mathbf{A}\mathbf{u}_{k+1}-\mathbf{v}\|_2^2 \leq \frac{1}{2}\|\mathbf{A}\mathbf{u}_k-\mathbf{v}\|_2^2+\tilde{C} \leq \cdots \leq \frac{1}{2}\|\mathbf{A}\mathbf{u}_1-\mathbf{v}\|_2^2+k\tilde{C}. \tag{A.3}$$

For the same reason, using (A.2) and (A.3) we get:

$$\|\mathbf{t}_{k+1}-\mathbf{L}_1\mathbf{u}_{k+1}\|_2 \leq \sqrt{\frac{1}{\rho_k^{\mathbf{t}}}}\|\mathbf{A}\mathbf{u}_1-\mathbf{v}\|_2+\sqrt{\frac{2\tilde{C}k}{\rho_k^{\mathbf{t}}}}, \tag{A.4}$$

$$\|\mathbf{z}_{k+1}-\mathbf{L}_2\mathbf{u}_{k+1}\|_2 \leq \sqrt{\frac{1}{\rho_k^{\mathbf{z}}}}\|\mathbf{A}\mathbf{u}_1-\mathbf{v}\|_2+\sqrt{\frac{2\tilde{C}k}{\rho_k^{\mathbf{z}}}}. \tag{A.5}$$

We now prove that the sequences $(\mathbf{t}_k)_{k=1}^{+\infty}$ and $(\mathbf{z}_k)_{k=1}^{+\infty}$ are Cauchy sequences. Starting from the expressions of $\mathbf{t}_{k+1}$ and $\mathbf{z}_{k+1}$ in Algorithm 1, applying the definition of bounded denoiser and the estimates (A.4) and (A.5) the following inequalities hold:

$$\begin{aligned}\|\mathbf{t}_{k+1}-\mathbf{t}_k\|_2 &\leq \|\mathcal{D}_{\sigma_k}^{\text{ext}}(\mathbf{L}_1\mathbf{u}_k)-\mathbf{L}_1\mathbf{u}_k\|_2+\|\mathbf{L}_1\mathbf{u}_k-\mathbf{t}_k\|_2 \leq \\ &\leq \sqrt{\frac{\alpha}{\rho_k^{\mathbf{t}}}}\sqrt{C_{\mathcal{D}^{\text{ext}}}}+\sqrt{\frac{1}{\rho_{k-1}^{\mathbf{t}}}}\|\mathbf{A}\mathbf{u}_1-\mathbf{v}\|_2+\sqrt{\frac{2\tilde{C}(k-1)}{\rho_{k-1}^{\mathbf{t}}}}\end{aligned} \tag{A.6}$$

$$\begin{aligned}\|\mathbf{z}_{k+1}-\mathbf{z}_k\|_2 &\leq \|\mathcal{D}_{\gamma_k}^{\text{int}}(\mathbf{L}_2\mathbf{u}_k)-\mathbf{L}_2\mathbf{u}_k\|_2+\|\mathbf{L}_2\mathbf{u}_k-\mathbf{z}_k\|_2 \leq \\ &\leq \sqrt{\frac{\beta}{\rho_k^{\mathbf{z}}}}\sqrt{C_{\mathcal{D}^{\text{int}}}}+\sqrt{\frac{1}{\rho_{k-1}^{\mathbf{z}}}}\|\mathbf{A}\mathbf{u}_1-\mathbf{v}\|_2+\sqrt{\frac{2\tilde{C}(k-1)}{\rho_{k-1}^{\mathbf{z}}}}.\end{aligned} \tag{A.7}$$

By assumption 3 $(\mathbf{z}_k)_{k=1}^{+\infty}$ and $(\mathbf{t}_k)_{k=1}^{+\infty}$ are Cauchy sequences. Hence, there exist $\mathbf{t}^*$ and $\mathbf{z}^*$ such that $\mathbf{t}_k \to \mathbf{t}^*$ and $\mathbf{z}_k \to \mathbf{z}^*$.

Furthermore, the following inequalities (which use (A.4) and (A.5), respectively) state that $\mathbf{L}_1\mathbf{u}_{k+1} \to \mathbf{t}^*$ and $\mathbf{L}_2\mathbf{u}_{k+1} \to \mathbf{z}^*$:

$$\|\mathbf{L}_1\mathbf{u}_{k+1}-\mathbf{t}^*\|_2 \leq \|\mathbf{L}_1\mathbf{u}_{k+1}-\mathbf{t}_{k+1}\|_2+\|\mathbf{t}_{k+1}-\mathbf{t}^*\|_2, \tag{A.8}$$

$$\|\mathbf{L}_2\mathbf{u}_{k+1}-\mathbf{z}^*\|_2 \leq \|\mathbf{L}_2\mathbf{u}_{k+1}-\mathbf{z}_{k+1}\|_2+\|\mathbf{z}_{k+1}-\mathbf{z}^*\|_2. \tag{A.9}$$

Now, we prove the convergence of the sequence $(\mathbf{u}_k)_{k=1}^{\infty}$ computed as in Algorithm 1. At step k, $\mathbf{u}_{k+1}$ is the solution of the convex minimization problem (8), therefore the first order optimality conditions lead:

$$\left(\frac{1}{\rho_k^{\mathbf{t}}}\mathbf{A}^T\mathbf{A} + \mathbf{L}_1^T\mathbf{L}_1 + \frac{\rho_k^{\mathbf{z}}}{\rho_k^{\mathbf{t}}}\mathbf{L}_2^T\mathbf{L}_2\right)\mathbf{u}_{k+1} = \frac{1}{\rho_k^{\mathbf{t}}}\mathbf{A}^T\mathbf{v} + \mathbf{L}_1^T\mathbf{t}_{k+1} + \frac{\rho_k^{\mathbf{z}}}{\rho_k^{\mathbf{t}}}\mathbf{L}_2^T\mathbf{z}_{k+1}. \quad \text{(A.10)}$$

If we define $\mathbf{M}_k := \frac{1}{\rho_k^{\mathbf{t}}}\mathbf{A}^T\mathbf{A} + \mathbf{L}_1^T\mathbf{L}_1 + \frac{\rho_k^{\mathbf{z}}}{\rho_k^{\mathbf{t}}}\mathbf{L}_2^T\mathbf{L}_2$, then $\forall\ k > 1$, $\mathbf{M}_k$ is invertible for assumption 2. Hence, we can write for each k:

$$\mathbf{u}_{k+1} = \mathbf{M}_k^{-1}\left(\frac{1}{\rho_k^{\mathbf{t}}}\mathbf{A}^T\mathbf{v} + \mathbf{L}_1^T\mathbf{t}_{k+1} + \frac{\rho_k^{\mathbf{z}}}{\rho_k^{\mathbf{t}}}\mathbf{L}_2^T\mathbf{z}_{k+1}\right). \quad \text{(A.11)}$$

We observe that the two sequences in the right hand side of (A.11), represented by $(\mathbf{M}_k^{-1})_{k=1}^{\infty}$ and by the term in parenthesis, are convergent pointwise (by assumption 3 and by considering the convergence of the sequences $(\mathbf{t}_k)_{k=1}^{\infty}$ and $(\mathbf{z}_k)_{k=1}^{\infty}$). By denoting as $\mathbf{u}^*$ the product of the two limits, we have proved that $\mathbf{u}_k \to \mathbf{u}^*$.
This concludes the proof. $\square$

We point out that this general proof applies also to the algorithm proposed in [15], for which no convergence results can be found in the literature. Moreover, we believe that with a small effort, our convergence result dealing with multiple denoisers can be extended to ADMM.

The fixed-point convergence Theorem A.1 entails that the iterations enter in a steady-state and does not guarantee that the fixed-point $\mathbf{u}^*$ is a minimum of an implicit defined regularized objective as in (3). However, in the experimental part, we have shown that the reached fixed-point $\mathbf{u}^*$ is a very good approximation of the desired image $\mathbf{u}$.

References

[1] C. Luo, Q. Liu, S. Gong, C. Yang, X. Hu, K. Nie, T. Niu, Cone-Beam Computed Tomography Deblurring Using an Overrelaxed Chambolle-pock Algorithm, IEEE Access 7 (2018) 1247–1259.

[2] P. Gravel, G. Beaudoin, J. A. De Guise, A method for modeling noise in medical images, IEEE Transactions on medical imaging 23 (10) (2004) 1221–1232.

[3] H. Lu, T. Hsiao, X. Li, Z. Liang, Noise properties of low-dose CT projections and noise treatment by scale transformations, in: 2001 IEEE Nuclear Science Symposium Conference Record (Cat. No. 01CH37310), Vol. 3, IEEE, 2001, pp. 1662–1666.

[4] Z. Al-Ameen, G. Sulong, Deblurring computed tomography medical images using a novel amended landweber algorithm, Interdisciplinary Sciences: Computational Life Sciences 7 (3) (2015) 319–325.

[5] Z. Al-Ameen, G. Sulong, M. Gapar, M. Johar, Reducing the gaussian blur artifact from ct medical images by employing a combination of sharpening filters and iterative deblurring algorithms, Journal of Theoretical and Applied Information Technology 46 (1) (2012) 31–36.

[6] M. Jiang, G. Wang, M. W. Skinner, J. T. Rubinstein, M. W. Vannier, Blind deblurring of spiral ct images, IEEE Transactions on medical imaging 22 (7) (2003) 837–845.

[7] D. Yim, B. Kim, S. Lee, A deep convolutional neural network for simultaneous denoising and deblurring in computed tomography, Journal of Instrumentation 15 (12) (2020) P12001.

[8] M. Bertero, Introduction to inverse problems in imaging, CRC press, 2020.

[9] L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D: nonlinear phenomena 60 (1-4) (1992) 259–268.

[10] B. Li, D. Que, Medical images denoising based on total variation algorithm, Procedia Environmental Sciences 8 (2011) 227–234.

[11] G. Landi, E. L. Piccolomini, An efficient method for nonnegatively constrained Total Variation-based denoising of medical images corrupted by poisson noise, Computerized Medical Imaging and Graphics 36 (1) (2012) 38–46.

[12] E. Loli Piccolomini, E. Morotti, A model-based optimization framework for iterative digital breast tomosynthesis image reconstruction, Journal of Imaging 7 (2). doi:10.3390/jimaging7020036.
URL https://www.mdpi.com/2313-433X/7/2/36

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

[14] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, C. A. Bouman, Plug-and-play priors for bright field electron tomography and sparse interpolation, IEEE Transactions on Computational Imaging 2 (4) (2016) 408–423.

[15] K. Zhang, W. Zuo, S. Gu, L. Zhang, Learning deep CNN denoiser prior for image restoration, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 3929–3938.

[16] U. S. Kamilov, H. Mansour, B. Wohlberg, A plug-and-play priors approach for solving nonlinear imaging inverse problems, IEEE Signal Processing Letters 24 (12) (2017) 1872–1876.

[17] R. Ahmad, C. A. Bouman, G. T. Buzzard, S. Chan, S. Liu, E. T. Reehorst, P. Schniter, Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery, IEEE signal processing magazine 37 (1) (2020) 105–116.

[18] T. Chen, X. Chen, W. Chen, H. Heaton, J. Liu, Z. Wang, W. Yin, Learning to Optimize: A Primer and A Benchmark (2021). `arXiv:2103.12828`.

[19] I. Mosseri, M. Zontak, M. Irani, Combining the power of internal and external denoising, in: IEEE international conference on computational photography (ICCP), IEEE, 2013, pp. 1–9.

[20] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Transactions on image processing 16 (8) (2007) 2080–2095.

[21] A. Buades, B. Coll, J.-M. Morel, A non-local algorithm for image denoising, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2, IEEE, 2005, pp. 60–65.

[22] D. Zoran, Y. Weiss, From learning models of natural image patches to whole image restoration, in: 2011 International Conference on Computer Vision, IEEE, 2011, pp. 479–486.

[23] Y. Chen, T. Pock, Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration, IEEE transactions on pattern analysis and machine intelligence 39 (6) (2016) 1256–1272.

[24] J. Xie, L. Xu, E. Chen, Image denoising and inpainting with deep neural networks, in: Advances in neural information processing systems, 2012, pp. 341–349.

[25] H. C. Burger, C. J. Schuler, S. Harmeling, Image denoising: Can plain neural networks compete with BM3D?, in: 2012 IEEE conference on computer vision and pattern recognition, IEEE, 2012, pp. 2392–2399.

[26] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, R. Timofte, Plug-and-Play image restoration with Deep Denoiser Prior, arXiv preprint.

[27] T. Meinhardt, M. Moller, C. Hazirbas, D. Cremers, Learning proximal operators: Using denoising networks for regularizing inverse imaging problems, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 1781–1790.

[28] A. Rond, R. Giryes, M. Elad, Poisson inverse problems by the plug-and-play scheme, Journal of Visual Communication and Image Representation 41 (2016) 96–108.

[29] L. He, Y. Wang, S. Gao, A support-denoiser-driven framework for single image restoration, Journal of Computational and Applied Mathematics (2021) 113495.

[30] V. Coli, E. L. Piccolomini, E. Morotti, L. Zanni, A fast gradient projection method for 3d image reconstruction from limited tomographic data, in: Journal of Physics: Conference Series, Vol. 904, IOP Publishing, 2017, p. 012013.

[31] K. Wei, A. Aviles-Rivero, J. Liang, Y. Fu, C.-B. Schönlieb, H. Huang, Tuning-free plug-and-play proximal algorithm for inverse imaging problems, in: International Conference on Machine Learning, PMLR, 2020, pp. 10158–10169.

[32] J. He, Y. Yang, Y. Wang, D. Zeng, Z. Bian, H. Zhang, J. Sun, Z. Xu, J. Ma, Optimizing a parameterized plug-and-play admm for iterative

low-dose ct reconstruction, IEEE transactions on medical imaging 38 (2) (2018) 371–382.

[33] D. Geman, C. Yang, Nonlinear image recovery with half-quadratic regularization, IEEE transactions on Image Processing 4 (7) (1995) 932–946.

[34] Y. Wang, J. Yang, W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM Journal on Imaging Sciences 1 (3) (2008) 248–272.

[35] P. C. Hansen, J. G. Nagy, D. P. Leary, R. Miller, Deblurring images: Matrices, spectra and filtering, Journal of Electronic Imaging 17 (1) (2008) 019901–019901.

[36] F. Yu, V. Koltun, Multi-scale context aggregation by dilated convolutions, arXiv preprint arXiv:1511.07122.

[37] D. Martin, C. Fowlkes, D. Tal, J. Malik, A Database of Human Segmented Natural Images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Proc. 8th Int'l Conf. Computer Vision, Vol. 2, 2001, pp. 416–423.

[38] V. A. Morozov, Methods for solving incorrectly posed problems, Springer, 1984.

[39] A. Hore, D. Ziou, Image quality metrics: Psnr vs. ssim, in: 2010 20th international conference on pattern recognition, IEEE, 2010, pp. 2366–2369.

[40] S. H. Chan, X. Wang, O. A. Elgendy, Plug-and-play ADMM for image restoration: Fixed-point convergence and applications, IEEE Transactions on Computational Imaging 3 (1) (2016) 84–98.