International Journal on Magnetic Particle Imaging
Vol 7, No 1, Article ID 2103001, 19 Pages

Infinite Science Publishing

Research Article

Deep image prior for 3D magnetic particle imaging: A quantitative comparison of regularization techniques on Open MPI dataset

Sören Dittmer[a,*] · Tobias Kluth[a] · Mads Thorstein Roar Henriksen[b] · Peter Maass[a]

[a] Center for Industrial Mathematics, University of Bremen, Bremen, Germany
[b] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark
*Corresponding author, email: sdittmer@math.uni-bremen.de

Received 07 August 2020; Accepted 02 February 2021; Published online 25 March 2021

Abstract

Magnetic particle imaging (MPI) is an imaging modality exploiting the nonlinear magnetization behavior of (super-) paramagnetic nanoparticles to obtain a space- and often also time-dependent concentration of a tracer consisting of these nanoparticles. MPI has a continuously increasing number of potential medical applications. One prerequisite for successful performance in these applications is a proper solution to the image reconstruction problem. More classical methods from inverse problems theory, as well as novel approaches from the field of machine learning, have the potential to deliver high-quality reconstructions in MPI. We investigate a novel reconstruction approach based on a deep image prior, which builds on representing the solution by a deep neural network. Novel approaches, as well as variational and iterative regularization techniques, are compared quantitatively in terms of peak signal-to-noise ratios and structural similarity indices on the publicly available `Open MPI dataset`.

I. Introduction

The imaging modality magnetic particle imaging (MPI) was invented by Gleich and Weizenecker in 2005 [1]. The goal of the technique is to reconstruct a concentration of (super-)paramagnetic iron oxide nanoparticles by applying a dynamic magnetic field. MPI benefits from high temporal resolution and potentially high spatial resolution which make it suitable for several *in-vivo* applications like imaging blood flow [2], long-term monitoring by utilizing a circulating tracer [3], flow estimation [4], tracking/guiding medical instruments [5], cancer detection [6], and cancer treatment by hyperthermia [7]. Moreover, the list of potential medical applications exploiting MPI is still growing, e.g., the recent prototype development of a human-sized scanner suitable for scanning a human's head [8] illustrates the potential of MPI for cerebral applications like stroke detection and monitoring [9]. For further technical background information on MPI, we refer to [10–12].

A critical prerequisite for potential medical applications is a proper solution to the image reconstruction problem in MPI, which aims for the determination of the space- and in some instances time-dependent tracer concentration. Neglecting concentration-dependent effects [13], likely to be caused by particle-particle interactions, the reconstruction is a linear ill-posed inverse problem [14], which for a given system matrix is typically solved by applying Tikhonov regularization in combination with the algebraic reconstruction technique and a nonneg-

ativity constraint [2, 15]. An alternative reconstruction approach for MPI is a direct inversion method exploiting Cartesian excitation patterns, the so-called x-space reconstruction [16, 17]. More generally, classical reconstruction methods from inverse problems theory taking into account further prior information, e.g., fused lasso regularization and directional total variation, have been applied to experimental [18] and simulated data [19]. More sophisticated approaches explicitly taking into account operator uncertainty have been proposed by using a total-least-squares approach combined with standard Tikhonov regularization as well as a sparsity-promoting penalty term [20]. Further efforts to improve the reconstruction in terms of accuracy and efficiency have been made, e.g., low-rank approximations and whitening [21] motivated by the non-polynomial decay of singular values [14], problem reduction via postprocessing [22], introducing an ℓ^1-data fidelity term [23], and many more (see [11] for an extensive overview).

Solving inverse problems by using techniques from the field of machine learning has received increased interest during the last years, and the number of their successful applications is continuously increasing [24]. One specific class of methods, based on the training of deep neural network architectures, is often summarized by the term deep learning-based methods. Specific examples for such image reconstruction methods include fully learned approaches [25, 26], unrolled iterative algorithms [27, 28], learned penalty terms [29], or postprocessing of classical reconstructions [30–32]. All of these methods have in common that they rely on training data. Often this training data is required in the form of tuples consisting of a given ground truth reconstruction and a corresponding measurement. In general, and also for MPI, this reliance on data leads to a chicken-and-egg type problem, since one relies on the assumption that sufficient ground truth data is available. In contrast, a deep image prior (DIP) approach [33] exploiting generative neural networks has been proposed recently for solving inverse problems in general. The DIP [34] is a novel regularization technique based on untrained neural network architectures. The basic idea is to use a neural network architecture/setting, that implicitly encodes a prior which encourages plausible reconstructions and relies on one single measurement only.

In the context of reconstruction methods for MPI, the investigation in the present paper is twofold, i.e.,

1. we introduce a three-dimensional deep image prior (DIP) for MPI, and
2. we compare various numerical reconstruction methods quantitatively as well as qualitatively in the context of the `Open MPI dataset` [35].

The subsequent part of the paper is structured as follows: In Section II, we provide a description of the used system matrix approach in MPI, variational reconstruction methods, a description of DIP, the experimental data set, and a brief description of the compared methods. In Section III, we continue with numerical results, including a quantitative comparison between DIP and classical methods, including variational and iterative approaches to image reconstruction. We conclude with a discussion in Section IV.

II. Methods

II.I. Magnetic particle imaging and preprocessing chain for derivation of the linear system of equations

In the following, we sketch the system matrix approach commonly used in MPI, see also [11] for further reading. We begin with some basic notation. Let $\Omega \subset \mathbb{R}^3$ be a bounded domain. Further, let $T > 0$ denote the maximal data acquisition time and $I := (0, T)$ the time interval during which the measurement process takes place.

The measured voltage signals $v_\ell : I \to \mathbb{R}$, $\ell = 1, \dots, L$, obtained at $L \in \mathbb{N}$ receive coil units, is given by a superposition of a signal $v_{\mathrm{P},\ell}$ caused by the particles and the direct feedthrough $v_{0,\ell}$ (background signal) caused mainly by the applied magnetic field. The inverse problem is thus to find the concentration $c : \Omega \to \mathbb{R}^+ \cup \{0\}$ from $\{v_\ell\}_{\ell=1}^L$:

$$v_\ell(t) = \int_\Omega c(x) s_\ell(x,t) \mathrm{d}x + v_{0,\ell}(t) = S_\ell c(t) + v_{0,\ell}(t), \quad \text{(II.1)}$$

where $S_\ell : L^2(\Omega) \to L^2(I)$ is the forward operator and where $s_\ell \in L^2(\Omega \times I)$ is the background-corrected system function.

The calibration procedure obtains single measurements from a small "delta" sample at predefined positions $\{x^{(i)}\}_{i=1,\dots,N} \in \Omega^N$ which builds the basis for the commonly used system matrix approach. For this purpose let $\Gamma \subset \mathbb{R}^3$ be a reference volume placed at the origin. The concentration phantoms are given by $c^{(i)} = c_0 \chi_{x^{(i)}+\Gamma}$ for some reference concentration $c_0 > 0$. Typical choices for Γ are small cubes. If $\{x^{(i)}+\Gamma\}_{i=1,\dots,N}$ form a partition of the domain Ω, the background-corrected measurements $v_\ell^{(i)} - v_{0,\ell}^{(i)} = S_\ell c^{(i)}$, $i = 1, \dots, N$, can then be used directly to characterize the system matrix S for L receive coil units ($v_{0,\ell}^{(i)}$, $\ell = 1, \dots, L$, are background measurements used for system matrix correction). For given phantom measurements v_ℓ, $\ell = 1, \dots, L$, we build the measurement vector v analogously. Both are then given by

$$S = \frac{1}{c_0} \begin{bmatrix} \mathrm{Re}((\langle v_1^{(i)} - v_{0,1}^{(i)}, \psi_j \rangle)_{j \in J_1, i=1,\dots,N}) \\ \mathrm{Im}((\langle v_1^{(i)} - v_{0,1}^{(i)}, \psi_j \rangle)_{j \in J_1, i=1,\dots,N}) \\ \vdots \\ \mathrm{Re}((\langle v_L^{(i)} - v_{0,L}^{(i)}, \psi_j \rangle)_{j \in J_L, i=1,\dots,N}) \\ \mathrm{Im}((\langle v_L^{(i)} - v_{0,L}^{(i)}, \psi_j \rangle)_{j \in J_L, i=1,\dots,N}) \end{bmatrix} \in \mathbb{R}^{M \times N}, \quad \text{(II.2)}$$

$$u = \begin{bmatrix} \frac{\mathrm{Re}((\langle u_1, \psi_j\rangle)_{j\in J_1})}{\mathrm{Im}((\langle u_1, \psi_j\rangle)_{j\in J_1})} \\ \vdots \\ \frac{\mathrm{Re}((\langle u_L, \psi_j\rangle)_{j\in J_L})}{\mathrm{Im}((\langle u_L, \psi_j\rangle)_{j\in J_L})} \end{bmatrix} \in \mathbb{R}^M, \qquad \text{(II.3)}$$

where $\{\psi_j\}_{j\in\mathbb{Z}}$ is the Fourier basis of time-periodic signals of $L^2(I)$, i.e., $\psi_j(t) = 1/\sqrt{T}(-1)^j e^{i2\pi jt/T}$, $j \in \mathbb{Z}$. For the purpose of preprocessing prior to reconstruction the sets $J_\ell \subset \mathbb{Z}$, $\ell = 1,\dots,L$ are restrictions to certain frequency indices, which also yield $M = 2\sum_\ell^L |J_\ell|$. Two frequency-selection approaches which are commonly combined result in the index sets J_ℓ, $\ell = 1,\dots,L$: a bandpass approach and SNR-type thresholding with threshold $\tau \geq 0$ (see [21, Sec. 2.2] for a more detailed description).

Let u_0 be the analogous measurement vector of the direct feedthrough, i.e., of an empty scanner. Then one obtains a measured signal from the L receive coils by $u^\delta = u + u_0 + \eta$ with noise vector η, $\|\eta\| \leq \delta$.

We thus obtain a linear system of equations $Sc = u^\delta - u_0$. Also, we include two processing steps, whitening, and low-rank approximation, which have been successfully used to improve reconstruction quality and computation times in MPI [21]. The linear equation system is multiplied with a whitening matrix W obtained from the diagonal covariance matrix of multiple background measurements. This allows to take into account the varying noise characteristic in the frequencies. For the low-rank approximation step, let $(\tilde{U}_K, \tilde{\Sigma}_K, \tilde{V}_K)$ be the randomized singular value decomposition (rSVD) for the $K \leq \min(M, N)$ largest singular values of the matrix WS. The main purpose of this step is dimension reduction instead of introducing an additional regularization in the problem setup. The severe ill-posedness of the imaging problem [14] and earlier investigations of the singular value energy ratio [21, Fig. 6] on the dataset used in the present work predict a reasonable dimension reduction without significant influence of truncation of the singular values.

Finally, this leaves us with the linear system

$$Ac = y^\delta \qquad \text{(II.4)}$$

with the processed matrix $A = \tilde{U}_K^t WS \in \mathbb{R}^{K\times N}$ and measurements $y^\delta = \tilde{U}_K^t W(u - u_0) \in \mathbb{R}^K$ ($\cdot^t$ denotes the transpose matrix).

In summary, we apply the following preprocessing steps to derive the final system of equations in (II.4) for a given SNR threshold τ (note that this is the only parameter in the preprocessing chain which is varied in the subsequent results section):

1. frequency selection by bandpass filtering and SNR-type thresholding, which removes too noisy rows in the system matrix,
2. concatenation of multiple receive coil units and splitting real and imaginary part,
3. weighting by diagonal whitening matrix, which takes into account varying standard deviations across the rows of the system matrix, and
4. system reduction by projection on subspace corresponding to K largest singular values of the whitened system matrix.

Note that steps 3 and 4 are motivated by improved reconstruction performance reported in the literature [21] but they are less standard in common MPI reconstruction approaches.

II.II. Classical reconstruction methods

Before describing the DIP method in the next subsection, we give a brief description of "classical reconstruction methods" to address the MPI problem which is given by an ill-posed operator equation ("classical" is to be understood in contrast to emerging reconstruction methods from the field of machine/deep learning).

Given some suitable spaces X and Y commonly assumed to be general Banach or Hilbert spaces, the reconstruction task is to compute a concentration $c \in X$ that agrees with a noisy measurement $y^\delta \in Y$ obtained by the linear measurement operator $A : X \to Y$, i.e., we want to find a "plausible" c such that

$$Ac \approx y^\delta. \qquad \text{(II.5)}$$

In line with the MPI setup specified in Section II.I, we consider the Hilbert spaces $X = \mathbb{R}^N$ and $Y = \mathbb{R}^K$ in the present paper.

For many inverse problems and also in MPI one follows a variational approach by minimizing a Tikhonov-type functional, i.e., a superposition of data fidelity and additive penalty term, which for the finite-dimensional system in the present work is given by

$$J_\lambda(c) = \frac{1}{p}\|Ac - y^\delta\|_p^p + \lambda R(c), \qquad \text{(II.6)}$$

where the regularization parameter $\lambda \geq 0$ and penalty term $R : \mathbb{R}^N \to \mathbb{R}_+$ include prior knowledge on the solution to obtain a stable reconstruction from noisy measurements. Common values for p are 1 and 2. Well-studied choices for R are, for example, $\frac{1}{2}\|\cdot\|_2^2$, $\|\cdot\|_1$, and total-variation terms TV [36]. The functionals are then minimized using suitable optimization techniques like, projected gradient descent type methods or incremental gradient descent method – which also include the Kaczmarz-type methods [37, 38] preferably used in MPI. In the latter case, one needs to be aware of the iterative nature of the algorithms, which can introduce an additional kind of regularization if a small iteration number is chosen not being sufficient for reaching convergence (early stopping).

10.18416/ijmpi.2021.2103001

II.III. Deep image prior

We now discuss the idea behind the deep image prior (DIP). The concept of a DIP was first introduced in [34], and it has been quickly adopted by fields like compressed sensing [39] and inverse problems [33]. Compared to classical reconstruction techniques, the DIP approach has shown improved reconstruction quality in image denoising [34], computed tomography [40], and positron emission tomography [41]. The core idea is to use the architecture of an untrained neural network $\varphi_\theta : \mathbb{R}^Z \to \mathbb{R}^N$ as a regularization for the reconstruction and to determine the network parameters $\theta \in \mathbb{R}^Q$ for a random but fixed input $z \in \mathbb{R}^Z$. More formally, in DIP one tries to minimize the data fidelity

$$J_{\mathrm{DIP}}(\theta) = \|A\varphi_\theta(z) - y^\delta\|_p^p, \tag{II.7}$$

with respect to θ, where usually $p = 2$, but throughout this paper we will use $p = 1$. The minimization is done iteratively by Adam [42]. This is in contrast to common approaches relying on the minimization of a Tikhonov-type functional as given in (II.6). One crucial prerequisite for the DIP is the proper choice of network architecture and often an adequate criterion for early stopping of the optimization. While the original DIP uses an architecture for 2D images, we use an architecture suitable for the 3D MPI problem.

As our regularizing architecture, we use an autoencoder. It is based on the original DIP paper's "skip-architecture"[34] and has approximately 3 million parameters. This means the network, like other DIPs, is highly overparameterized. This overparameterization seems necessary for tasks where a large amount of information is missing – as the paper [43] demonstrates for inpainting tasks. In detail, our architecture differs in the following aspects from the "skip-architecture":

- Like [40], in a two-dimensional setting, we find that removing the skip-connections leads to improved reconstructions. Therefore we use the architecture without the skip-connections, i.e., we reduce it to an autoencoder.
- Naturally, we replaced the two-dimensional convolutional layers by three-dimensional ones.
- Our final activation function is ReLU, not Sigmoid. We use ReLU because we know that our particle concentration is non-negative, but do not know an upper bound.

We depict the architecture in Figure 6 of the Appendix.

We conclude this subsection with a discussion of the relationship between the DIP and the variational approach, i.e., *rewriting variational regularizations as deep image priors*: As demonstrated in [33], the DIP can be related to variational approaches to inverse problems. For these theoretical considerations we have to assume optimization up to convergence (no early stopping) and continuous differentiability of the network with respect to its parameters θ. Then one can apply the idea of Lagrange multipliers [44] to the constrained optimization problem

$$\min_c \|Ac - y^\delta\|_2^2 \text{ s.t. } \|c - \varphi_\theta(z)\|_2^2 = 0. \tag{II.8}$$

We would like to point out that it is also possible to write most of the common regularization functionals R in the context of the variational problem

$$\min_c \|Ac - y^\delta\|_2^2 + \lambda R(c) \tag{II.9}$$

as a DIP – although this can require exotic architectures. One way to show this is to assume continuous differentiability of R and utilize Lagrange multipliers to rewrite the Expression (II.9) as the constrained optimization problem

$$\min_c \|Ac - y^\delta\|_2^2 \text{ s.t. } R(c) = \gamma(\lambda). \tag{II.10}$$

If we further assume the existence of an R specific functional $g_\gamma : \mathrm{domain}(R) = \mathbb{R}^Q = \mathbb{R}^N \to \mathbb{R}$ s.t. $R(g_\gamma(\theta)\theta) = \gamma$, we can rewrite (II.10) in terms of the problem

$$\min_\theta \|Ag_\gamma(\theta)\theta - y^\delta\|_2^2. \tag{II.11}$$

The solution θ of this problem yields the solution to the original problem in (II.9) via $g_\gamma(\theta)\theta$. The assumption that such an g_γ exists is fulfilled by most common penalty terms in use, like total-variation [36], and ℓ^p-norms which fulfil this property in the form of positive homogeneity. E.g., for a penalty term of the form

$$R(c) = \|c\|_p^p,$$

and $c \neq 0$, we have the functional $g_\gamma(c) = \frac{\sqrt[p]{\gamma}}{\|c\|_p}$ and therefore the DIP

$$\varphi_\theta(z) = \frac{\sqrt[p]{\gamma}}{\|\theta\|_p}\theta. \tag{II.12}$$

II.IV. Experimental data and image quality assessment

For the experimental evaluation we use the 3D `open MPI dataset` [35] (downloaded from `https://www.tuhh.de/ibi/research/open-mpi-data.html`, last accessed on April 02, 2020) provided in the MPI Data Format (MDF) [45]. The system matrix is measured using a cuboid sample of size 2 mm × 2 mm × 1 mm and a 3D Lissajous-type FFP excitation while obtaining averaged empty scanner measurements every 19 calibration scans. The calibration is carried out with a Perimag tracer having a concentration of 100 mmol/l. The field-of-view has a size of 38 mm × 38 mm × 19 mm, and the sample positions have a distance of 2 mm in x- and y-direction and 1 mm in z-direction, resulting in $19 \times 19 \times 19 = 6859 = N$ voxels.

Measurements are averaged over 1000 repetitions, and with each phantom, an empty measurement with 1000 repetitions is provided. By taking the mean of the empty measurements provided with the phantom, we obtain $v_{0,\ell}$, $\ell = 1,\dots,L$, respectively v_0, in Section II.I. For the background removal in the system matrix we used $v_{0,\ell}^{(i)} = v_{0,\ell}$ for any $i = 1,\dots,N$ and, if not mentioned differently, use $K = 2000$ for the low-rank approximation step discussed in Section II.I.

We use the "shape" and "resolution" phantoms provided in the dataset. We state the detailed description for the sake of completeness. It can also be found on at `https://www.tuhh.de/ibi/research/open-mpi-data.html` or in [23]. The "shape" phantom is a cone defined by a 1 mm radius tip, an apex angle of 10 degrees, and a height of 22 mm. The total volume is 683.9 μl. Perimag tracer with a concentration of 50 mmol/l is used. See [23, Figure 3] for a schematic illustration, where the plots are adapted from the `Open MPI dataset`. The "resolution" phantom consists of 5 tubes filled with Perimag tracer with a concentration of 50 mmol/l. The 5 tubes have a common origin on one side of the phantom, and extend in different angles from the origin within the x-y- and y-z-planes. In the z-direction, the angles in the y-z-plane are chosen smaller (10 deg and 15 deg) than in x-y-plane (20 deg and 30 deg); see [23, Figure 4] for the illustration.

For image quality assessment, we exploit the approach in [23], where the authors extracted voxel images from the CAD drawings and included position uncertainty in peak-signal-to-noise-ratio (PSNR) and structural similarity measure (SSIM) [46]. More precisely, a reference image $c_0 : \mathbb{R}^3 \to \mathbb{R}_+$ is used in combination with shifts $\Delta r \in \mathscr{R}$, where $\mathscr{R}$ is the set of all possible position shifts in the neighborhood $[-3\text{mm}, 3\text{mm}]^3$ with a step size 0.5mm in each direction ($|\mathscr{R}| = 2197$). Then the L^2-scalar products of $c_0(r + \Delta r)$ and the piecewise constant basis functions $\{\psi_j\}_j$ with respect to the voxel grid yields the reference image $x_{\text{ref},\Delta r} \in \mathbb{R}^N$. Following the approach in [23] we define the following conservative image quality measures

$$\epsilon_{\text{PSNR}}(x) = \max_{\Delta r \in \mathscr{R}} \text{PSNR}(x, x_{\text{ref},\Delta r}) \tag{II.13}$$

$$\epsilon_{\text{SSIM}}(x) = \max_{\Delta r \in \mathscr{R}} \text{SSIM}(x, x_{\text{ref},\Delta r}) \tag{II.14}$$

which are used for quantitative comparison in the following. Here, we use a data range of 100 for the SSIM computations.

II.V. Reconstruction methods used for comparison

In the subsequent quantitative and qualitative comparison, we distinguish three classes of reconstruction methods, which we evaluate on the two phantoms of the 3D `Open MPI dataset` described in Section II.IV:

1. **DIP**: Reconstructions based on the DIP approach minimizing a data fidelity term as described in Section II.III using Adam with varying numbers of iterations.
2. **VAR**: Reconstructions based on the minimization of a Tikhonov-type functional consisting of a data-fidelity and a penalty term as outlined in Section II.II. Minimization is performed by AMSGrad [47] until convergence is reached such that results rely on the *variational* regularization only.
3. **KACZ**: Analogous to VAR, but the minimization realized by a *Kaczmarz*-type algorithm with varying numbers of iterations. Effectively it yields reconstructions, which are results of a hybrid regularization, mixing iterative and variational elements.

We encode our methods based on the class, the respective data fidelity, and the penalty terms, i.e., `Method class D data fidelity + P penalty`. All methods rely on a certain number of parameters controlling the degree of regularization. We performed a discrete optimization with respect to these parameters to obtain optimal results in terms of PSNR and SSIM. We thus compare the best possible performance of the different methods. From an application point of view, without knowledge about ground truth an appropriate regularization parameter choice rule is required, which is an important and related direction of research but beyond the scope of this article. The methods of interest in this comparison are specified in more detail in the following list:

- **DIP $\mathbf{D}\ell^1 + \mathbf{P}$-**: We minimize the functional (II.7) with $p = 1$ only as it led to superior reconstruction performance when compared to the case $p = 2$.
  - Our fixed random input to the network, z, has entries coming from the uniform distribution between 0 and 0.7 and has, like the output, the shape $(1, 19, 19, 19)$ where 1 is the number of channels and 19 is the edge length of the concentration cube we aim to reconstruct.
  - The encoder steps down-sample by a factor of 2 and have 64, 128 and 256 channels respectively. The decoder is symmetric to this.
  - We minimize the functional for up to 20000 iterations with Adam [42] and for the different learning rates $\alpha_i = 10^{-i}$, for $i = 3, 4, 5$, and the standard momenta settings $\beta = (0.9, 0.999)$.

  For the purpose of parameter optimization (with regard to image quality of the reconstruction) we extracted reconstructions after iterations $s \in I_{\text{DIP}} := \{1, 2, \dots, 10, 12, \dots, 30, 35, \dots, 50, 60, \dots, 150, 175, \dots, 500, 600, \dots, 2000, 2500, \dots, 5000, 6000, \dots, 20000\}$ to exploit early stopping.
- **KACZ $\mathbf{D}\ell^2$ + $\mathbf{P}\ell^2$**: This is one of the commonly used techniques in MPI which minimizes the functional $\frac{1}{2}\|Ac - y^\delta\|_2^2 + \rho\frac{1}{2}\|c\|_2^2$ with non-negativity

constraints by a row action method for inconsistent systems of equations [48] (see [38, Sec. 4] or [21, Alg. 1] for the detailed algorithm). For the parameter optimization $\rho \in \{0.5^{i-1} | i = 1, \cdots 40\} =: \mathscr{P}$ were used. Like all the following KACZ reconstructions, we consider all reconstructions after the same sweeps/iterations s as for the DIP up to a maximum of 500 sweeps (i.e., going 500 times over all rows in the system matrix) to choose the best reconstruction with respect to image quality, i.e., $s \in I_{\text{KACZ}} := I_{\text{DIP}} \cap \{1, 2, \ldots, 500\}$.

- **KACZ Dℓ^2 + P(ℓ^1+ℓ^2)**: This is an extension of the previous algorithm for inconsistent systems of equations taking into account an additional sparsity-promoting ℓ^1-penalty term when minimizing the functional $\|Ac - y^\delta\|_2^2 + \rho \frac{1}{2}\|c\|_2^2 + \lambda\|c\|_1$ with non-negativity constraints. We exploit a splitting method [49, Sec. 9.4.1] first applying one sweep over the matrix as in the previous method followed by applying a soft shrinkage operator before starting a new sweep over the matrix. A similar approach applying the soft shrinkage operator in each row action can be found in [50] for consistent systems and an explicit description can be found in [51, Alg.1]. Both regularization parameters are taken from $\mathscr{P}$ as above (i.e., resulting in 40^2 time $|I_{\text{KACZ}}|$ parameter combinations).
- **KACZ Dℓ^2 + Pℓ^1**: We consider the previous functional for $\rho = 0$, i.e., $\frac{1}{2}\|Ac - y^\delta\|_2^2 + \lambda_i\|c\|_1$ with non-negativity constraints via the previous algorithm (see also [52, Alg. 2] for an example implementation). We ran reconstructions for all parameter choices $\lambda \in \mathscr{P}$ times the number of iteration numbers contained in I_{KACZ}.
- **KACZ+TSVD Dℓ^2 + Pℓ^1**: The normalization step in Kaczmarz divides by the sum of the current row's squared ℓ^2-norm and ρ. When using $\rho = 0$ – like in the previous method – and rows with small norms exist, this leads to instabilities. Therefore, in the presence of low norm rows, it makes sense to remove them when one applies Kaczmarz with no ℓ^2-regularization. Due to the low-rank approximation in preprocessing and the problem's ill-posedness, the matrix contains rows close to zero. We thus also used only the 32, 64, 128, 256, 512, and 1024 rows with the largest norms in the previous method (and corresponding entries of the measurement). I.e., we set K in Section II.I to these values, which is equivalent to a truncated SVD (TSVD). Reconstructions for $(\lambda, s) \in \mathscr{P} \times I_{\text{KACZ}}$ are computed.
- **VAR**: Here, we minimize a functional of the form II.6, where we chose the penalty parameter as above and the penalty function as ℓ^1, ℓ^2 and TV. In the results section the parameter λ belongs to ℓ^1, ρ to ℓ^2, and μ to TV. We also compare $p = 1$ and $p = 2$. Minimization of the functional is performed up to convergence by applying AMSGrad [47] with a learning rate of 10^{-2} and again $\beta = (0.9, 0.999)$ (`pyTorch` [53] standard settings) for 500 iterations, after each iteration we project c to be non-negative.

All methods are implemented in `python` [54] mainly using the packages `numpy` [55] and `pyTorch` [53].

III. Results

First, we compare the results quantitatively based on the optimal PSNR and SSIM values, which are presented in Tables 1 and 2, for the shape and resolution phantom, respectively. The corresponding optimal parameters that led to these values can be found in the supplementary material in Tables 5 and 6. Note that all the data discussed in this section relates to the whitened case discussed in the preprocessing part of Section II.I. The supplementary material contains the analogous parts of the non-whitened case, which qualitatively leads to similar results, though, with PSNRs being on average 0.5 lower and SSIMs being on average 0.12 lower for peak values.

Qualitatively we present the results in Figures 1, 2, and 3. Figure 1 utilizes a three-dimensional visualization to compare the best PSNR reconstruction with the ground truth phantom for the shape and resolution phantom respectively. Both best reconstructions were achieved by our DIP approach. We display the loss, PSNRs, and SSIMs over time of these two reconstructions in the Figures 7 and 8 of the Appendix. Figures 2 and 3 display two-dimensional cross sections of the reconstructions corresponding to the values of Tables 1 and 2 – except for VAR Dℓ^2 since these methods result in the worst reconstructions. The main observations are as follows:

1. The proposed DIP approach achieves the overall highest PSNR and SSIM values for both the shape and the resolution phantom. Even for SNR-type threshold settings (τ's/columns), in which the DIP produces inferior peak results, they are usually comparable to the best method. We found that for the DIP, the PSNR rapidly rose within the first 50-100 iterations. It always reached 95% of the run's top PSNR within the first 150-1500 iterations, usually even over 98%. On a GeForce GTX 1080 Ti, this equates to 3.2-32 seconds for $\tau = 3$. For the shape phantom, the top PSNR was typically achieved after the first few hundred iterations, after which the PSNR began to drop significantly. For the resolution phantom, we also saw a rapid early rise in PSNR, which turned into a prolonged slow increase over several thousand iterations. An equally slow decline followed this slow increase. Consequently, the peak PSNR for the resolution phantom occurred typically quite late.

10.18416/ijmpi.2021.2103001

Methods	PSNRs				SSIMs			
⟋	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
DIP $D\ell^1+$ P-	28.70	29.36	**29.81**	**<u>29.94</u>**	**0.968**	**0.968**	**<u>0.973</u>**	**<u>0.973</u>**
KACZ $D\ell^2+P\ell^2$	29.12	29.31	28.89	28.79	0.957	0.958	0.955	0.953
KACZ $D\ell^2+P\ell^1$	26.21	26.63	18.01	17.09	0.886	0.888	0.534	0.534
KACZ+TSVD $D\ell^2+P\ell^1$	28.84	29.62	28.76	28.53	0.953	0.953	0.953	0.947
KACZ $D\ell^2+P(\ell^1+\ell^2)$	**29.52**	**29.70**	29.43	29.15	0.963	0.964	0.961	0.958
VAR $D\ell^1+P\ell^1$	21.45	21.58	21.45	22.64	0.856	0.865	0.840	0.876
VAR $D\ell^1+P\ell^2$	27.57	27.89	27.50	27.81	0.932	0.910	0.876	0.885
VAR $D\ell^1+$ P TV	25.98	25.84	26.29	27.31	0.932	0.917	0.930	0.938
VAR $D\ell^2+P\ell^1$	17.70	21.28	23.03	24.06	0.646	0.807	0.882	0.897
VAR $D\ell^2+P\ell^2$	20.14	24.53	26.11	26.33	0.684	0.854	0.913	0.906
VAR $D\ell^2+$ P TV	18.96	22.78	25.75	25.58	0.648	0.818	0.915	0.914

Table 1: PSNR and SSIM values as achieved by the different SNR-type thresholds τ settings for the shape phantom. Largest values for each column are in bold font. The overall largest value is underlined. The corresponding parameter choices including the iteration number s for DIP and KACZ can be found in Table 5.

Methods	PSNRs				SSIMs			
⟋	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
DIP $D\ell^1+$ P-	31.73	32.46	**<u>32.92</u>**	**32.33**	0.945	**<u>0.952</u>**	**0.951**	0.940
KACZ $D\ell^2+P\ell^2$	31.58	32.08	31.89	31.71	0.946	0.947	0.949	0.945
KACZ $D\ell^2+P\ell^1$	31.82	32.08	30.24	29.39	0.943	0.944	0.887	0.829
KACZ+TSVD $D\ell^2+P\ell^1$	**31.94**	**32.57**	31.80	31.35	0.947	0.948	0.947	0.942
KACZ $D\ell^2+P(\ell^1+\ell^2)$	31.82	32.38	32.10	31.76	**0.949**	0.949	0.949	**0.946**
VAR $D\ell^1+P\ell^1$	30.43	30.82	29.91	30.25	0.915	0.919	0.901	0.904
VAR $D\ell^1+P\ell^2$	31.42	31.59	31.42	31.14	0.945	0.947	0.942	0.936
VAR $D\ell^1+$ P TV	31.22	31.61	31.47	31.24	0.940	0.948	0.933	0.928
VAR $D\ell^2+P\ell^1$	29.00	28.93	30.08	30.93	0.803	0.860	0.910	0.905
VAR $D\ell^2+P\ell^2$	29.51	30.23	30.88	31.24	0.839	0.889	0.932	0.931
VAR $D\ell^2+$ P TV	29.26	29.54	30.44	30.79	0.823	0.878	0.925	0.925

Table 2: PSNR and SSIM values as achieved by the different SNR-type thresholds τ settings for the resolution phantom. Largest values for each column are in bold font. The overall largest value is underlined. The corresponding parameter choices including the iteration number s for DIP and KACZ can be found in Table 6.

2. Quantitatively, the VAR $D\ell^1$ reconstructions are superior to the VAR $D\ell^2$ reconstructions as the Tables 1, and 2 show. Although in the present work, we project the data as well as the noise onto the subspace spanned by the singular functions corresponding to the largest singular values, this finding is in line with the observation in [23]. On a GeForce GTX 1080 Ti, one run of such a VAR method took about 1.6 seconds for $\tau=3$.

3. Within the KACZ methods, we can observe that the combination of ℓ^2- and ℓ^1-term is superior to the standard KACZ method solely using the ℓ^2-term.

4. The KACZ methods, which utilize early-stopping (see s-values in Tables 5 and 6), produce better results than the purely variational methods (VAR). On a Xeon E5-2687W v4, one sweep of such a KACZ method took approximately 0.3 seconds for $\tau=3$, opposed to around 0.6 seconds on a GeForce GTX 1080 Ti. The best reconstructions required 1-7 sweeps. You can find the exact number of sweeps for the different reconstructions in the Appendix's parameter tables.

5. The KACZ $D\ell^2+P\ell^1$ method can suffer from instability issues if it has to deal with matrix rows of small norm. However, further reducing the rank using TSVD (method KACZ+TSVD $D\ell^2+P\ell^1$) has a beneficial influence on the reconstruction quality, but in almost all cases, it cannot reach the performance of KACZ $D\ell^2+P(\ell^1+\ell^2)$ which is an alternative approach to stabilize the reconstruction.

6. Particularly for the resolution phantom, we can observe that in almost all cases, larger SNR-type thresholds result in worse reconstructions, which is not a surprise as SNR-thresholding is known to remove phantom information if the threshold is chosen too large [21, 22].

10.18416/ijmpi.2021.2103001

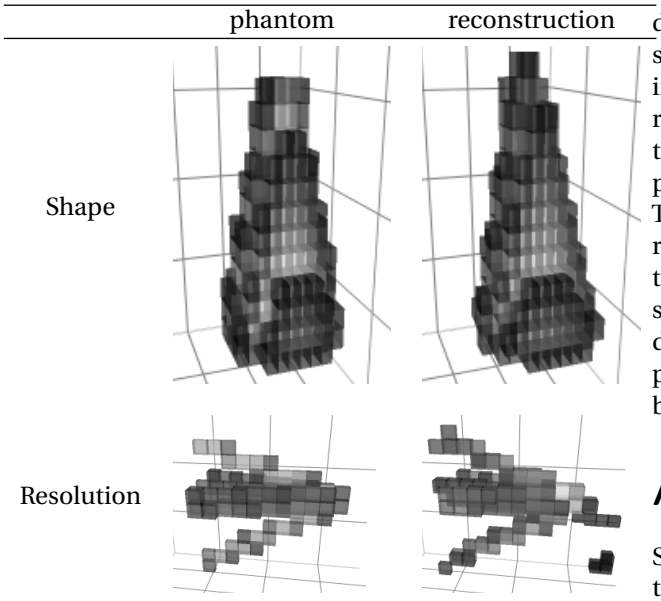


Figure 1: Qualitative comparison of the phantoms and the best reconstructions in terms of PSNR (all achieved by the proposed DIP). The gray scale goes from 0 (black) to 60 or 40 (white), for the shape and resolution phantom respectively.

IV. Discussion and conclusion

In summary, we have multiple main findings illustrated by the quantitative results in the present work. First, the proposed deep image prior approach based on an autoencoder architecture outperforms all other methods considered in this study. The superior behavior of the DIP is particularly found for smaller numbers of training steps (early stopping) when determining the network parameters. Second, the results for the KACZ class, which is also characterized by limited numbers of iterations (early stopping), illustrate the success of the algebraic reconstruction technique with sole ℓ^2-penalty in MPI [11]. The data and noise structure in MPI seem to be beneficial for the iterative nature of DIP and KACZ methods. Within the KACZ methods, it turned out that the combination of ℓ^2- and ℓ^1-terms in the penalty can result in improved reconstructions compared to the commonly used reconstruction method in MPI. One drawback of the superior reachable image quality in this method is the additional regularization parameter, which is often difficult to choose. An automated choice of these parameters is highly desirable for one-click image reconstructions in applications. The intense study of regularization parameter choice rules for all considered methods in the context of MPI is beyond the scope of the present work and thus remains future work. Third, this study emphasizes the need for a careful discussion and distinction between different regularization techniques. Variational approaches (VAR) strongly rely on a proper choice of the data fidelity term, i.e., prior knowledge on the data space and noise distribution. A standard ℓ^2-data fidelity term is not well suited for MPI reconstructions. Similar to the findings in [23], an ℓ^1-data fidelity term is better suited for MPI reconstruction, not only in the DIP method but also in the VAR methods. This holds even if the data space is projected onto a subspace spanned by singular vectors. The latter finding also emphasizes the need for future research on the identification and proper treatment of the MPI noise structure for the purpose of image reconstruction. Future work also include a validation on larger datasets and a comparison to learned methods [24]. In particular the latter requires a sufficiently large MPI data basis which needs to be set up in the future.

Acknowledgments

S. Dittmer and T. Kluth acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - project number 281474342/GRK2224/1 “Pi^3 : Parameter Identification - Analysis, Algorithms, Applications” and T. Kluth further acknowledges support by the project “MPI^2” funded by the Federal Ministry of Education and Research (BMBF, project no. 05M16LBA).

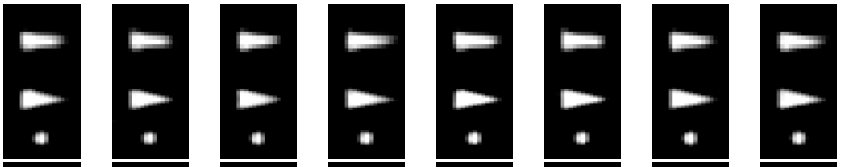
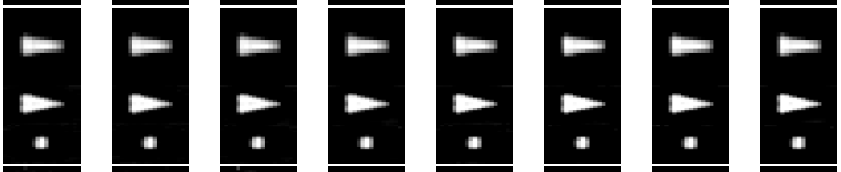
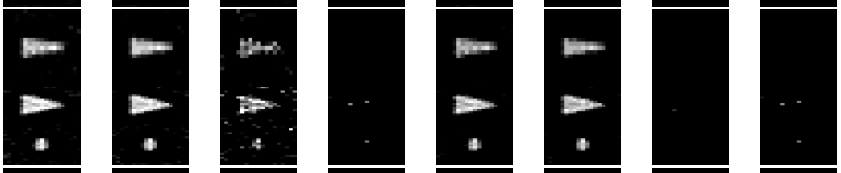
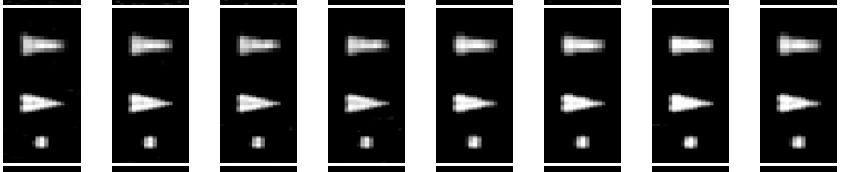
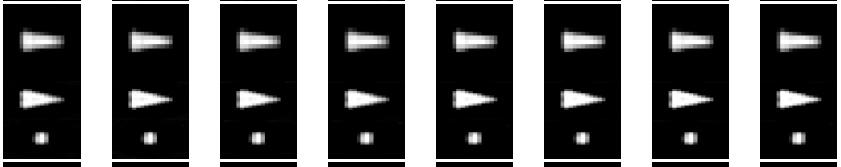
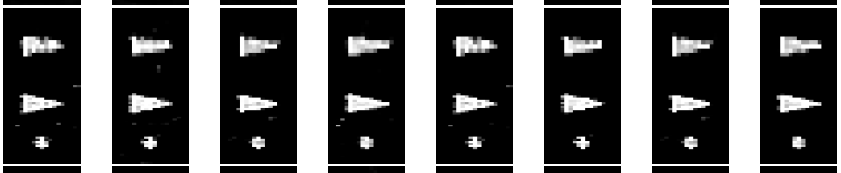
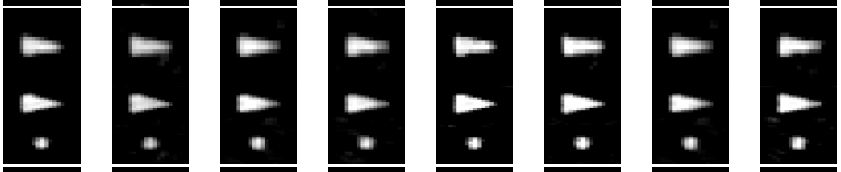
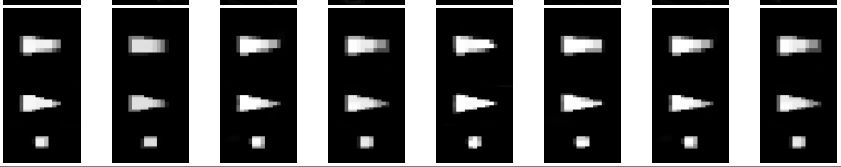
Methods	PSNRs				SSIMs			
	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$
DIP $D\ell^1 + P$-								
KACZ $D\ell^2 + P\ell^2$								
KACZ $D\ell^2 + P\ell^1$								
KACZ+TSVD $D\ell^2 + P\ell^1$								
KACZ $D\ell^2 + P(\ell^1+\ell^2)$								
VAR $D\ell^1 + P\ell^1$								
VAR $D\ell^1 + P\ell^2$								
VAR $D\ell^1 + P$ TV								

Figure 2: The best reconstructions for the shape phantom, corresponding to the values in Tables 1 and 5 for varying SNR-type thresholds τ. The color scale goes from 0 (black) to 60 (white). Each image is separated into three vertically stacked parts, which are separated by red lines. Each of these parts represents one of the three central slices/planes of the three-dimensional reconstruction (x-y-, x-z-, and y-z-plane from top to bottom).

10.18416/ijmpi.2021.2103001

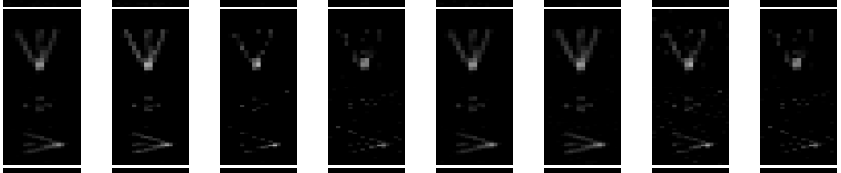
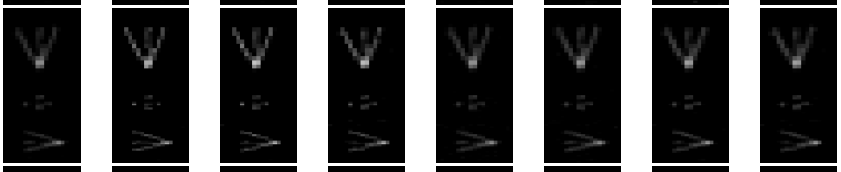
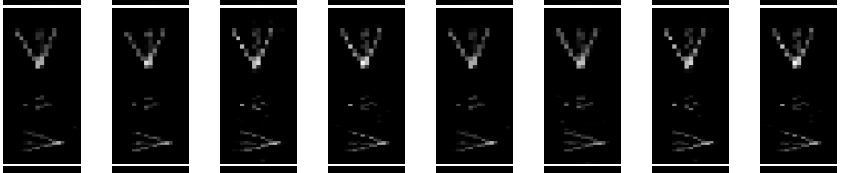
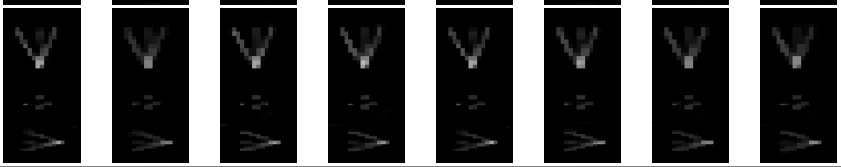
Methods	PSNRs				SSIMs			
	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$
DIP $D\ell^1 + P$-								
KACZ $D\ell^2 + P\ell^2$								
KACZ $D\ell^2 + P\ell^1$								
KACZ+TSVD $D\ell^2 + P\ell^1$								
KACZ $D\ell^2 + P(\ell^1+\ell^2)$								
VAR $D\ell^1 + P\ell^1$								
VAR $D\ell^1 + P\ell^2$								
VAR $D\ell^1 + P$ TV								

Figure 3: The best reconstructions for the resolution phantom, corresponding to the values in Tables 2 and 6 for varying SNR-type thresholds τ. The color scale goes from 0 (black) to 40 (white). Each image is separated into three vertically stacked parts, which are separated by red lines. Each of these parts represents one of the three central slices/planes of the three-dimensional reconstruction (x-y-, x-z-, and y-z-plane from top to bottom).

10.18416/ijmpi.2021.2103001

References

[1] B. Gleich and J. Weizenecker. Tomographic imaging using the nonlinear response of magnetic particles. *Nature*, 435(7046):1214–1217, 2005, doi:10.1038/nature03808.

[2] J. Weizenecker, B. Gleich, J. Rahmer, H. Dahnke, and J. Borgert. Three-dimensional real-time in vivo magnetic particle imaging. *Physics in Medicine and Biology*, 54(5):L1–L10, 2009, doi:10.1088/0031-9155/54/5/L01.

[3] A. P. Khandhar, P. Keselman, S. J. Kemp, R. M. Ferguson, P. W. Goodwill, S. M. Conolly, and K. M. Krishnan. Evaluation of PEG-coated iron oxide nanoparticles as blood pool tracers for preclinical magnetic particle imaging. *Nanoscale*, 9(3):1299–1306, 2017, doi:10.1039/C6NR08468K.

[4] J. Franke, R. Lacroix, H. Lehr, M. Heidenreich, U. Heinen, and V. Schulz. MPI Flow Analysis Toolbox exploiting pulsed tracer information - an aneurysm phantom proof. *International Journal on Magnetic Particle Imaging*, 3(1), 2017, doi:10.18416/IJMPI.2017.1703020.

[5] J. Haegele, J. Rahmer, B. Gleich, J. Borgert, H. Wojtczyk, N. Panagiotopoulos, T. M. Buzug, J. Barkhausen, and F. M. Vogt. Magnetic Particle Imaging: Visualization of Instruments for Cardiovascular Intervention. *Radiology*, 265(3):933–938, 2012, doi:10.1148/radiol.12120424.

[6] E. Y. Yu, M. Bishop, B. Zheng, R. M. Ferguson, A. P. Khandhar, S. J. Kemp, K. M. Krishnan, P. W. Goodwill, and S. M. Conolly. Magnetic Particle Imaging: A Novel in Vivo Imaging Platform for Cancer Detection. *Nano Letters*, 17(3):1648–1654, 2017, doi:10.1021/acs.nanolett.6b04865.

[7] K. Murase, M. Aoki, N. Banura, K. Nishimoto, A. Mimura, T. Kuboyabu, and I. Yabata. Usefulness of Magnetic Particle Imaging for Predicting the Therapeutic Effect of Magnetic Hyperthermia. *Open Journal of Medical Imaging*, 05(02):85–99, 2015, doi:10.4236/ojmi.2015.52013.

[8] M. Graeser, F. Thieben, P. Szwargulski, F. Werner, N. Gdaniec, M. Boberg, F. Griese, M. Möddel, P. Ludewig, D. van de Ven, O. M. Weber, O. Woywode, B. Gleich, and T. Knopp. Human-sized magnetic particle imaging for brain applications. *Nature Communications*, 10(1):1936, 2019, doi:10.1038/s41467-019-09704-x.

[9] P. Ludewig, N. Gdaniec, J. Sedlacik, N. D. Forkert, P. Szwargulski, M. Graeser, G. Adam, M. G. Kaul, K. M. Krishnan, R. M. Ferguson, A. P. Khandhar, P. Walczak, J. Fiehler, G. Thomalla, C. Gerloff, T. Knopp, and T. Magnus. Magnetic Particle Imaging for Real-Time Perfusion Imaging in Acute Stroke. *ACS Nano*, 11(10):10480–10488, 2017, doi:10.1021/acsnano.7b05784.

[10] T. Knopp and T. M. Buzug, Magnetic Particle Imaging: An Introduction to Imaging Principles and Scanner Instrumentation. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, doi:10.1007/978-3-642-04199-0.

[11] T. Knopp, N. Gdaniec, and M. Möddel. Magnetic particle imaging: from proof of principle to preclinical applications. *Physics in Medicine & Biology*, 62(14):R124–R178, 2017, doi:10.1088/1361-6560/aa6c99.

[12] T. Kluth. Mathematical models for magnetic particle imaging. *Inverse Problems*, 34(8):083001, 2018, doi:10.1088/1361-6420/aac535.

[13] N. Löwa, P. Radon, O. Kosch, and F. Wiekhorst. Concentration Dependent MPI Tracer Performance. *International Journal on Magnetic Particle Imaging*, 2(1), 2016, doi:10.18416/IJMPI.2016.1601001.

[14] T. Kluth, B. Jin, and G. Li. On the degree of ill-posedness of multi-dimensional magnetic particle imaging. *Inverse Problems*, 34(9):095006, 2018, doi:10.1088/1361-6420/aad015.

[15] T. Knopp, J. Rahmer, T. F. Sattel, S. Biederer, J. Weizenecker, B. Gleich, J. Borgert, and T. M. Buzug. Weighted iterative reconstruction for magnetic particle imaging. *Physics in Medicine and Biology*, 55(6):1577–1589, 2010, doi:10.1088/0031-9155/55/6/003.

[16] P. W. Goodwill and S. M. Conolly. The X-Space Formulation of the Magnetic Particle Imaging Process: 1-D Signal, Resolution, Bandwidth, SNR, SAR, and Magnetostimulation. *IEEE Transactions on Medical Imaging*, 29(11):1851–1859, 2010, doi:10.1109/TMI.2010.2052284.

[17] P. W. Goodwill and S. M. Conolly. Multidimensional X-Space Magnetic Particle Imaging. *IEEE Transactions on Medical Imaging*, 30(9):1581–1590, 2011, doi:10.1109/TMI.2011.2125982.

[18] M. Storath, C. Brandt, M. Hofmann, T. Knopp, J. Salamon, A. Weber, and A. Weinmann. Edge Preserving and Noise Reducing Reconstruction for Magnetic Particle Imaging. *IEEE Transactions on Medical Imaging*, 36(1):74–85, 2017, doi:10.1109/TMI.2016.2593954.

[19] C. Bathke, T. Kluth, C. Brandt, and P. Maaß. Improved image reconstruction in magnetic particle imaging using structural a priori information. *International Journal on Magnetic Particle Imaging*, 3(1), 2017, doi:10.18416/IJMPI.2017.1703015.

[20] T. Kluth and P. Maass. Model uncertainty in magnetic particle imaging: Nonlinear problem formulation and model-based sparse reconstruction. *International Journal on Magnetic Particle Imaging*, 3(2), 2017, doi:10.18416/IJMPI.2017.1707004.

[21] T. Kluth and B. Jin. Enhanced reconstruction in magnetic particle imaging by whitening and randomized SVD approximation. *Physics in Medicine & Biology*, 64(12):125026, 2019, doi:10.1088/1361-6560/ab1a4f.

[22] T. Knopp and M. Hofmann. Online reconstruction of 3D magnetic particle imaging data. *Physics in Medicine and Biology*, 61(11):N257–N267, 2016, doi:10.1088/0031-9155/61/11/N257.

[23] T. Kluth and B. Jin. L1 data fitting for robust reconstruction in magnetic particle imaging: quantitative evaluation on Open MPI dataset. *International Journal on Magnetic Particle Imaging*, 6(2), 2020, doi:10.18416/IJMPI.2020.2012001.

[24] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019, doi:10.1017/S0962492919000059.

[25] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen. Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487–492, 2018, doi:10.1038/nature25988.

[26] J. He, Y. Wang, and J. Ma. Radon Inversion via Deep Learning. *IEEE Transactions on Medical Imaging*, 39(6):2076–2087, 2020, doi:10.1109/TMI.2020.2964266.

[27] J. Adler and O. Öktem. Solving ill-posed inverse problems using iterative deep neural networks. *Inverse Problems*, 33(12):124007, 2017, doi:10.1088/1361-6420/aa9581.

[28] J. Adler and O. Öktem. Learned Primal-Dual Reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1322–1332, 2018, doi:10.1109/TMI.2018.2799231.

[29] S. Lunz, O. Öktem, and C.-B. Schönlieb, Adversarial regularizers in inverse problems, in *NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 8516–8525, 2018.

[30] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang. Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network. *IEEE Transactions on Medical Imaging*, 36(12):2524–2535, 2017, doi:10.1109/TMI.2017.2715284.

[31] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017, doi:10.1109/TIP.2017.2713099.

[32] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang. Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. *IEEE Transactions on Medical Imaging*, 37(6):1348–1357, 2018, doi:10.1109/TMI.2018.2827462.

[33] S. Dittmer, T. Kluth, P. Maass, and D. Otero Baguer. Regularization by Architecture: A Deep Prior Approach for Inverse Problems. *Journal of Mathematical Imaging and Vision*, 62(3):456–470, 2020, doi:10.1007/s10851-019-00923-x.

[34] V. Lempitsky, A. Vedaldi, and D. Ulyanov, Deep Image Prior, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9446–9454, IEEE, 2018. doi:10.1109/CVPR.2018.00984.

[35] T. Knopp, P. Szwargulski, F. Griese, and M. Gräser. OpenMPIData: An initiative for freely accessible magnetic particle imaging data. *Data in Brief*, 28:104971, 2020, doi:10.1016/j.dib.2019.104971.

[36] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992, doi:10.1016/0167-2789(92)90242-F.

[37] S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin International de l'Académie Polonaise des Sciences et des Lettres*, (35)pp. 355–357, 1937.

[38] A. Dax. On Row Relaxation Methods for Large Constrained Least Squares Problems. *SIAM Journal on Scientific Computing*, 14(3):570–584, 1993, doi:10.1137/0914036.

[39] D. Van Veen, A. Jalal, M. Soltanolkotabi, E. Price, S. Vishwanath, and A. G. Dimakis. Compressed Sensing with Deep Image Prior and Learned Regularization, 2018. arXiv: 1806.06438. URL: http://arxiv.org/abs/1806.06438.

[40] D. O. Baguer, J. Leuschner, and M. Schmidt. Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems*, 36(9):094004, 2020, doi:10.1088/1361-6420/aba415.

[41] K. Gong, C. Catana, J. Qi, and Q. Li. PET Image Reconstruction Using Deep Image Prior. *IEEE Transactions on Medical Imaging*, 38(7):1655–1665, 2019, doi:10.1109/TMI.2018.2888491.

[42] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization, 2014. arXiv: 1412.6980. URL: http://arxiv.org/abs/1412.6980.

[43] R. Heckel and P. Hand. Deep Decoder: Concise Image Representations from Untrained Non-convolutional Networks, 2018. arXiv: 1810.03982. URL: http://arxiv.org/abs/1810.03982.

[44] D. P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods. Elsevier, 1982, doi:10.1016/C2013-0-10366-2.

[45] T. Knopp, T. Viereck, G. Bringout, M. Ahlborg, A. von Gladiss, C. Kaethner, A. Neumann, P. Vogel, J. Rahmer, and M. Möddel. MDF: Magnetic Particle Imaging Data Format, 2016. arXiv: 1602.06072. URL: http://arxiv.org/abs/1602.06072.

[46] A. Hore and D. Ziou, Image Quality Metrics: PSNR vs. SSIM, in *2010 20th International Conference on Pattern Recognition*, 2366–2369, IEEE, 2010. doi:10.1109/ICPR.2010.579.

[47] S. J. Reddi, S. Kale, and S. Kumar. On the Convergence of Adam and Beyond, 2019. arXiv: 1904.09237. URL: http://arxiv.org/abs/1904.09237.

[48] G. T. HERMAN, A. LENT, and H. HURWITZ. A Storage-Efficient Algorithm for Finding the Regularized Solution of a Large, Inconsistent System of Equations. *IMA Journal of Applied Mathematics*, 25(4):361–366, 1980, doi:10.1093/imamat/25.4.361.

[49] O. Scherzer, Handbook of Mathematical Methods in Imaging, O. Scherzer, Ed. New York, NY: Springer New York, 2011, doi:10.1007/978-0-387-92920-0.

[50] D. A. Lorenz, S. Wenger, F. Schopfer, and M. Magnor, A sparse Kaczmarz solver and a linearized Bregman method for online compressed sensing, in *2014 IEEE International Conference on Image Processing (ICIP)*, 1347–1351, IEEE, 2014. doi:10.1109/ICIP.2014.7025269.

[51] T. Kluth, C. Bathke, M. Jiang, and P. Maass. Joint super-resolution image reconstruction and parameter identification in imaging operator: Analysis of bilinear operator equations, numerical solution, and application to magnetic particle imaging, 2020, doi:10.1088/1361-6420/abc2fe.

[52] F. Lieb and T. Knopp. A Wavelet Based Sparse Row-Action Method for Image Reconstruction in Magnetic Particle Imaging, 2020. arXiv: 2003.13787. URL: http://arxiv.org/abs/2003.13787.

[53] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, Automatic differentiation in PyTorch. 2017,

[54] G. V. Rossum and F. L. Drake, Python 3 Reference Manual. CreateSpace, ISBN: 1441412697.

[55] T. E. Oliphant, A guide to NumPy. Trelgol Publishing, 2006,

Supplementary material of
Deep image prior for 3D magnetic particle imaging:
A quantitative comparison of regularization techniques on `Open MPI dataset`

10.18416/ijmpi.2021.2103001

Methods	PSNRs				SSIMs			
/	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
DIP Dℓ^1 + P-	26.70	**29.59**	**29.79**	**<u>29.81</u>**	0.926	**<u>0.966</u>**	**0.962**	**0.965**
KACZ Dℓ^2 + Pℓ^2	27.50	27.69	27.78	27.78	0.925	0.935	0.943	0.939
KACZ Dℓ^2 + Pℓ^1	24.06	24.03	16.71	14.69	0.794	0.786	0.458	0.209
KACZ+TSVD Dℓ^2 + Pℓ^1	27.14	27.23	27.39	27.54	0.919	0.916	0.937	0.939
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	**28.03**	28.28	28.36	28.20	**0.939**	0.943	0.947	0.944
VAR Dℓ^1 + Pℓ^1	18.25	19.93	21.47	21.83	0.749	0.799	0.809	0.842
VAR Dℓ^1 + Pℓ^2	23.78	25.84	24.88	25.44	0.835	0.834	0.792	0.798
VAR Dℓ^1 + P TV	24.37	26.21	24.87	26.42	0.848	0.916	0.866	0.900
VAR Dℓ^2 + Pℓ^1	17.04	19.01	23.62	24.96	0.596	0.684	0.887	0.896
VAR Dℓ^2 + Pℓ^2	19.85	21.32	25.12	25.43	0.638	0.700	0.887	0.883
VAR Dℓ^2 + P TV	19.52	20.69	24.70	24.78	0.613	0.685	0.887	0.891

Table 3: PSNR and SSIM values as achieved by the different SNR-type thresholds τ settings for the shape phantom <u>without</u> whitening. Largest values for each column are in bold font. The overall largest value is underlined. The corresponding parameter choices including the iteration number s for DIP and KACZ can be found in Table 7.

Methods	PSNRs				SSIMs			
/	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
DIP Dℓ^1 + P-	**<u>32.02</u>**	**31.97**	**31.66**	**31.63**	0.912	**0.945**	0.939	0.943
KACZ Dℓ^2 + Pℓ^2	31.16	31.38	31.45	31.27	0.938	0.941	0.946	0.944
KACZ Dℓ^2 + Pℓ^1	31.66	31.58	29.34	28.99	0.938	0.937	0.828	0.761
KACZ+TSVD Dℓ^2 + Pℓ^1	31.68	31.71	31.22	30.86	0.939	0.941	0.936	0.934
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	31.56	31.69	31.65	31.46	**0.942**	0.944	**<u>0.947</u>**	**0.945**
VAR Dℓ^1 + Pℓ^1	29.85	30.57	29.80	29.72	0.896	0.904	0.905	0.889
VAR Dℓ^1 + Pℓ^2	30.87	31.34	31.00	30.93	0.921	0.940	0.936	0.934
VAR Dℓ^1 + P TV	31.02	31.29	30.61	30.35	0.916	0.938	0.929	0.927
VAR Dℓ^2 + Pℓ^1	28.87	29.17	29.14	29.82	0.782	0.829	0.882	0.898
VAR Dℓ^2 + Pℓ^2	29.49	29.53	30.42	30.42	0.811	0.847	0.906	0.911
VAR Dℓ^2 + P TV	29.20	29.39	29.87	29.85	0.808	0.845	0.899	0.908

Table 4: PSNR and SSIM values as achieved by the different SNR-type thresholds τ settings for the resolution phantom <u>without</u> whitening. Largest values for each column are in bold font. The overall largest value is underlined. The corresponding parameter choices including the iteration number s for DIP and KACZ can be found in Table 8.

Methods	PSNRs				SSIMs			
	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$
DIP $D\ell^1 + P$-								
KACZ $D\ell^2 + P\ell^2$								
KACZ $D\ell^2 + P\ell^1$								
KACZ+TSVD $D\ell^2 + P\ell^1$								
KACZ $D\ell^2 + P(\ell^1+\ell^2)$								
VAR $D\ell^1 + P\ell^1$								
VAR $D\ell^1 + P\ell^2$								
VAR $D\ell^1 + P$ TV								

Figure 4: The best reconstructions for the shape phantom without whitening, corresponding to the values in Tables 3 and 7. The color scale goes from 0 (black) to 60 (white). Each image is separated into three vertically stacked parts, which are separated by red lines. Each of these parts represents one of the three central slices/planes of the three-dimensional reconstruction (x-y-, x-z-, and y-z-plane from top to bottom).

10.18416/ijmpi.2021.2103001

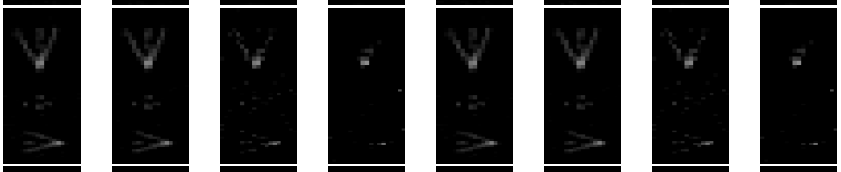
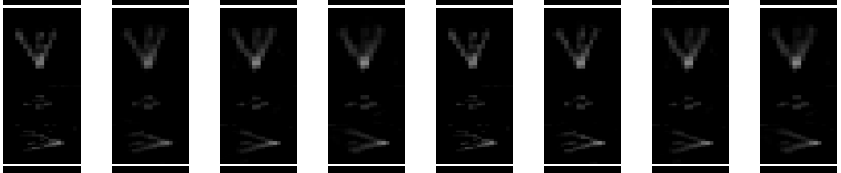
Methods	PSNRs				SSIMs			
	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$	$\tau = 0$	$\tau = 1$	$\tau = 3$	$\tau = 5$
DIP Dℓ^1 + P-								
KACZ Dℓ^2 + Pℓ^2								
KACZ Dℓ^2 + Pℓ^1								
KACZ+TSVD Dℓ^2 + Pℓ^1								
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)								
VAR Dℓ^1 + Pℓ^1								
VAR Dℓ^1 + Pℓ^2								
VAR Dℓ^1 + P TV								

Figure 5: The best reconstructions for the resolution phantom without whitening, corresponding to the values in Tables 4 and 8. The color scale goes from 0 (black) to 40 (white). Each image is separated into three vertically stacked parts, which are separated by red lines. Each of these parts represents one of the three central slices/planes of the three-dimensional reconstruction.

10.18416/ijmpi.2021.2103001

	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
	PSNRs			
DIP Dℓ^1 + P-	α_3, $s=200$	α_4, $s=150$	α_4, $s=225$	α_3, $s=350$
KACZ Dℓ^2 + Pℓ^2	ρ_{15}, $s=5$	ρ_{12}, $s=4$	ρ_{12}, $s=4$	ρ_{12}, $s=5$
KACZ Dℓ^2 + Pℓ^1	λ_5, $s=1$	λ_5, $s=1$	λ_3, $s=1$	λ_1, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=512$, λ_6, $s=1$	$d=512$, λ_6, $s=1$	$d=512$, λ_5, $s=1$	$d=512$, λ_5, $s=1$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{16}, λ_6, $s=2$	ρ_{14}, λ_6, $s=2$	ρ_{13}, λ_6, $s=2$	ρ_{13}, λ_6, $s=2$
VAR Dℓ^1 + Pℓ^1	$\lambda_1$6	λ_8	λ_6	λ_7
VAR Dℓ^1 + Pℓ^2	ρ_6	ρ_4	ρ_4	ρ_4
VAR Dℓ^1 + P TV	μ_8	μ_5	μ_7	μ_6
VAR Dℓ^2 + Pℓ^1	λ_{50}	λ_{50}	λ_{11}	λ_{12}
VAR Dℓ^2 + Pℓ^2	ρ_9	ρ_8	ρ_9	ρ_9
VAR Dℓ^2 + P TV	μ_{10}	μ_8	μ_{12}	μ_{13}
	SSIMs			
DIP Dℓ^1 + P-	α_3, $s=400$	α_4, $s=250$	α_3, $s=375$	α_3, $s=250$
KACZ Dℓ^2 + Pℓ^2	ρ_{15}, $s=10$	ρ_{12}, $s=11$	ρ_{11}, $s=14$	ρ_{12}, $s=12$
KACZ Dℓ^2 + Pℓ^1	λ_4, $s=1$	λ_4, $s=1$	λ_1, $s=1$	λ_1, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=256$, λ_6, $s=2$	$d=256$, λ_6, $s=2$	$d=128$, λ_7, $s=7$	$d=256$, λ_6, $s=2$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{16}, λ_7, $s=6$	ρ_{13}, λ_6, $s=4$	ρ_{13}, λ_6, $s=5$	ρ_{13}, λ_6, $s=5$
VAR Dℓ^1 + Pℓ^1	$\lambda_1$2	λ_6	λ_6	λ_6
VAR Dℓ^1 + Pℓ^2	ρ_7	ρ_5	ρ_4	ρ_5
VAR Dℓ^1 + P TV	μ_{10}	μ_7	μ_7	μ_6
VAR Dℓ^2 + Pℓ^1	λ_{50}	λ_{50}	λ_{14}	λ_{12}
VAR Dℓ^2 + Pℓ^2	ρ_{10}	ρ_8	ρ_{10}	ρ_{10}
VAR Dℓ^2 + P TV	μ_{13}	μ_{12}	μ_{12}	μ_{13}

Table 5: Optimal parameters creating the results for the shape phantom with whitening as displayed in Table 1. $\rho_i = 0.5^{i-1}$, λ_i, μ_i analogously and $\alpha_i = 10^{-i}$.

	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
	PSNRs			
DIP Dℓ^1 + P-	α_4, $s=1100$	α_3, $s=5000$	α_3, $s=11000$	α_4, $s=4000$
KACZ Dℓ^2 + Pℓ^2	ρ_{17}, $s=5$	ρ_{17}, $s=25$	ρ_{16}, $s=25$	ρ_{16}, $s=27$
KACZ Dℓ^2 + Pℓ^1	λ_7, $s=1$	λ_7, $s=4$	λ_5, $s=2$	λ_5, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=1024$, λ_7, $s=1$	$d=1024$, λ_7, $s=8$	$d=512$, λ_8, $s=14$	$d=512$, λ_8, $s=10$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{19}, λ_8, $s=2$	ρ_{19}, λ_9, $s=35$	ρ_{18}, λ_8, $s=18$	ρ_{18}, λ_9, $s=24$
VAR Dℓ^1 + Pℓ^1	λ_8	λ_5	λ_8	λ_7
VAR Dℓ^1 + Pℓ^2	ρ_5	ρ_3	ρ_2	ρ_4
VAR Dℓ^1 + P TV	μ_{11}	μ_8	μ_9	μ_9
VAR Dℓ^2 + Pℓ^1	λ_{13}	λ_{12}	λ_{14}	λ_{15}
VAR Dℓ^2 + Pℓ^2	ρ_{12}	ρ_{11}	ρ_{12}	ρ_{13}
VAR Dℓ^2 + P TV	μ_{15}	μ_{13}	μ_{16}	μ_{18}
	SSIMs			
DIP Dℓ^1 + P-	α_4, $s=3500$	α_4, $s=1900$	α_3, $s=3500$	α_4, $s=900$
KACZ Dℓ^2 + Pℓ^2	ρ_{18}, $s=11$	ρ_{14}, $s=11$	ρ_{14}, $s=11$	ρ_{14}, $s=12$
KACZ Dℓ^2 + Pℓ^1	λ_7, $s=1$	λ_8, $s=1$	λ_6, $s=1$	λ_5, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=1024$, λ_8, $s=1$	$d=1024$, λ_8, $s=1$	$d=512$, λ_8, $s=3$	$d=512$, λ_8, $s=3$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{19}, λ_9, $s=5$	ρ_{16}, λ_8, $s=3$	ρ_{15}, λ_9, $s=5$	ρ_{15}, λ_{10}, $s=8$
VAR Dℓ^1 + Pℓ^1	λ_8	λ_8	λ_7	λ_7
VAR Dℓ^1 + Pℓ^2	ρ_5	ρ_3	ρ_3	ρ_3
VAR Dℓ^1 + P TV	μ_{11}	μ_9	μ_8	μ_8
VAR Dℓ^2 + Pℓ^1	λ_{40}	λ_{17}	λ_{40}	λ_{15}
VAR Dℓ^2 + Pℓ^2	ρ_{12}	ρ_{11}	ρ_{12}	ρ_{12}
VAR Dℓ^2 + P TV	μ_{15}	μ_{16}	μ_{17}	μ_{17}

Table 6: Optimal parameters creating the results for the resolution phantom with whitening as displayed in Table 2. $\rho_i = 0.5^{i-1}$, λ_i, μ_i analogously and $\alpha_i = 10^{-i}$.

	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
	PSNRs			
DIP Dℓ^1 + P-	α_3, $s=325$	α_4, $s=350$	α_3, $s=350$	α_3, $s=325$
KACZ Dℓ^2 + Pℓ^2	ρ_{15}, $s=6$	ρ_{13}, $s=7$	ρ_{12}, $s=5$	ρ_{12}, $s=7$
KACZ Dℓ^2 + Pℓ^1	λ_4, $s=1$	λ_4, $s=1$	λ_1, $s=1$	λ_1, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=512$, λ_5, $s=1$	$d=512$, λ_5, $s=1$	$d=128$, λ_6, $s=4$	$d=128$, λ_6, $s=4$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{16}, λ_6, $s=3$	ρ_{15}, λ_6, $s=3$	ρ_{13}, λ_6, $s=3$	ρ_{14}, λ_6, $s=3$
VAR Dℓ^1 + Pℓ^1	λ_7	λ_8	λ_6	λ_7
VAR Dℓ^1 + Pℓ^2	ρ_7	ρ_6	ρ_5	ρ_6
VAR Dℓ^1 + P TV	μ_8	μ_8	μ_7	μ_8
VAR Dℓ^2 + Pℓ^1	λ_6	λ_{40}	λ_{11}	λ_{11}
VAR Dℓ^2 + Pℓ^2	ρ_9	ρ_8	ρ_{10}	ρ_{10}
VAR Dℓ^2 + P TV	μ_{10}	μ_9	μ_{12}	μ_{12}
	SSIMs			
DIP Dℓ^1 + P-	α_3, $s=475$	α_4, $s=400$	α_3, $s=425$	α_3, $s=700$
KACZ Dℓ^2 + Pℓ^2	ρ_{15}, $s=12$	ρ_{13}, $s=15$	ρ_{13}, $s=20$	ρ_{13}, $s=20$
KACZ Dℓ^2 + Pℓ^1	λ_4, $s=1$	λ_4, $s=1$	λ_1, $s=1$	λ_1, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=256$, λ_6, $s=2$	$d=256$, λ_6, $s=2$	$d=128$, λ_7, $s=7$	$d=128$, λ_7, $s=7$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{16}, λ_7, $s=8$	ρ_{15}, λ_7, $s=9$	ρ_{13}, λ_7, $s=12$	ρ_{14}, λ_7, $s=12$
VAR Dℓ^1 + Pℓ^1	λ_{27}	λ_7	λ_6	λ_6
VAR Dℓ^1 + Pℓ^2	ρ_8	ρ_6	ρ_6	ρ_6
VAR Dℓ^1 + P TV	μ_8	μ_8	μ_8	μ_8
VAR Dℓ^2 + Pℓ^1	λ_{40}	λ_{40}	λ_{40}	λ_{12}
VAR Dℓ^2 + Pℓ^2	ρ_{10}	ρ_{10}	ρ_{12}	ρ_{12}
VAR Dℓ^2 + P TV	μ_{11}	μ_{15}	μ_{16}	μ_{15}

Table 7: Optimal parameters creating the results for the shape phantom without whitening as displayed in Table 3. $\rho_i = 0.5^{i-1}$, λ_i, μ_i analogously and $\alpha_i = 10^{-i}$.

	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
	PSNRs			
DIP Dℓ^1 + P-	α_4, $s=14000$	α_4, $s=6000$	α_3, $s=3500$	α_4, $s=4000$
KACZ Dℓ^2 + Pℓ^2	ρ_{17}, $s=16$	ρ_{16}, $s=5$	ρ_{15}, $s=6$	ρ_{16}, $s=29$
KACZ Dℓ^2 + Pℓ^1	λ_7, $s=1$	λ_7, $s=1$	λ_5, $s=1$	λ_4, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=1024$, λ_7, $s=1$	$d=1024$, λ_8, $s=1$	$d=512$, λ_7, $s=1$	$d=512$, λ_7, $s=4$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{19}, λ_8, $s=2$	ρ_{18}, λ_8, $s=2$	ρ_{16}, λ_8, $s=2$	ρ_{17}, λ_8, $s=3$
VAR Dℓ^1 + Pℓ^1	λ_8	λ_7	λ_8	λ_8
VAR Dℓ^1 + Pℓ^2	ρ_6	ρ_4	ρ_4	ρ_4
VAR Dℓ^1 + P TV	μ_{11}	μ_{10}	μ_{11}	μ_{10}
VAR Dℓ^2 + Pℓ^1	λ_{18}	λ_{14}	λ_{15}	λ_{15}
VAR Dℓ^2 + Pℓ^2	ρ_{13}	ρ_{12}	ρ_{13}	ρ_{13}
VAR Dℓ^2 + P TV	μ_{15}	μ_{16}	μ_{16}	μ_{16}
	SSIMs			
DIP Dℓ^1 + P-	α_4, $s=4500$	α_4, $s=1600$	α_3, $s=1800$	α_4, $s=6000$
KACZ Dℓ^2 + Pℓ^2	ρ_{18}, $s=12$	ρ_{17}, $s=14$	ρ_{15}, $s=14$	ρ_{16}, $s=17$
KACZ Dℓ^2 + Pℓ^1	λ_7, $s=1$	λ_7, $s=1$	λ_5, $s=1$	λ_4, $s=1$
KACZ+TSVD Dℓ^2 + Pℓ^1	$d=1024$, λ_8, $s=1$	$d=1024$, λ_8, $s=2$	$d=512$, λ_8, $s=2$	$d=256$, λ_{10}, $s=12$
KACZ Dℓ^2 + P(ℓ^1+ℓ^2)	ρ_{20}, λ_8, $s=4$	ρ_{18}, λ_9, $s=5$	ρ_{17}, λ_9, $s=6$	ρ_{17}, λ_9, $s=8$
VAR Dℓ^1 + Pℓ^1	λ_{17}	λ_9	λ_9	λ_8
VAR Dℓ^1 + Pℓ^2	ρ_6	ρ_5	ρ_4	ρ_4
VAR Dℓ^1 + P TV	μ_{12}	μ_{10}	μ_9	μ_{10}
VAR Dℓ^2 + Pℓ^1	λ_{40}	λ_{18}	λ_{17}	λ_{40}
VAR Dℓ^2 + Pℓ^2	ρ_{12}	ρ_{12}	ρ_{13}	ρ_{14}
VAR Dℓ^2 + P TV	μ_{14}	μ_{17}	μ_{17}	μ_{19}

Table 8: Optimal parameters creating the results for the resolution phantom without whitening as displayed in Table 4. $\rho_i = 0.5^{i-1}$, λ_i, μ_i analogously and $\alpha_i = 10^{-i}$.

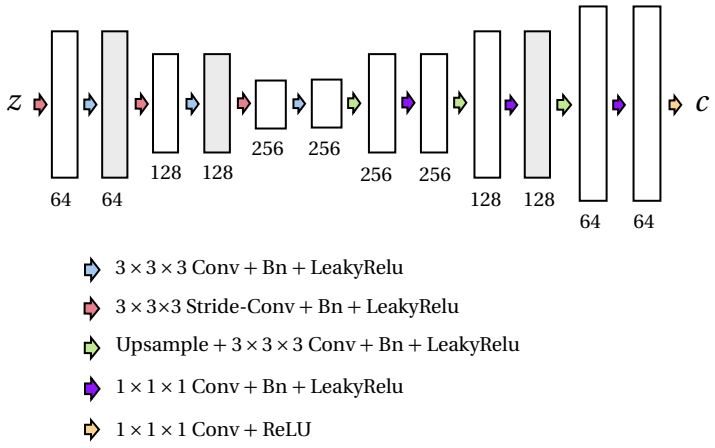


Figure 6: The autoencoder architecture of the 3D DIP. (The style of the depiction is following [40].)

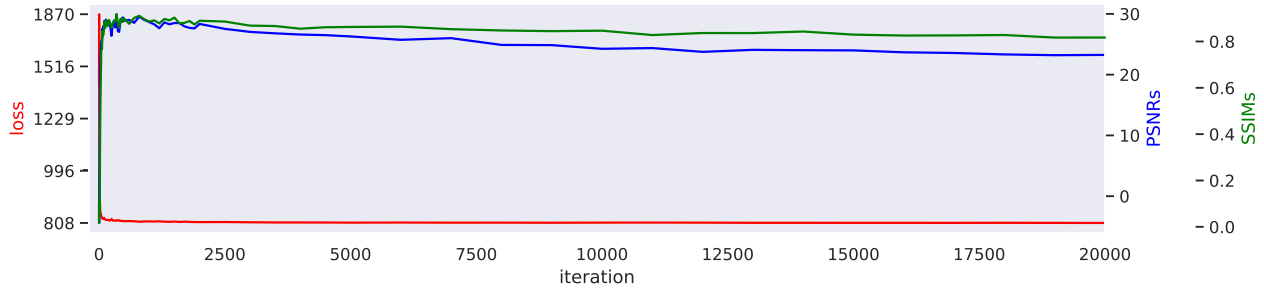


Figure 7: Plot of the loss, PSNRs, and SSIMs over time for the best (DIP) shape reconstruction.

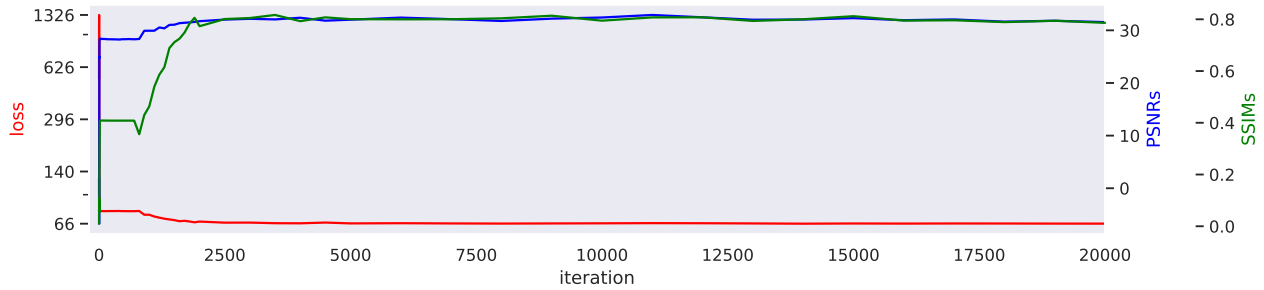


Figure 8: Plot of the loss, PSNRs, and SSIMs over time for the best (DIP) resolution reconstruction.

10.18416/ijmpi.2021.2103001