Deep Gauss–Newton for phase retrieval

Kannara Mom, Max Langer, Bruno Sixou

Deep Gauss-Newton for phase retrieval

KANNARA MOM[1,*], MAX LANGER[1,2], AND BRUNO SIXOU[1]

[1] Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1206, F-69621 Villeurbanne, France
[2] Currently at Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, F-38000 Grenoble, France
[*] kannara.mom@creatis.insa-lyon.fr

Compiled January 24, 2023

We propose the Deep Gauss-Newton (DGN) algorithm. The DGN allows to take into account the knowledge of the forward model in a deep neural network by unrolling a Gauss-Newton optimization method. No regularization or step size need to be chosen, they are learned through convolutional neural networks. The proposed algorithm does not require an initial reconstruction and is able to retrieve simultaneously the phase and absorption from a single-distance diffraction pattern. The DGN method was applied to both simulated and experimental data and permitted large improvements of the reconstruction error and of the resolution compared to a state of the art iterative method and another neural network based reconstruction algorithm.

http://dx.doi.org/10.1364/ao.XX.XXXXXX

Phase contrast imaging has higher sensitivity compared to attenuation-based techniques, and permits tomography at the microscopic and nanoscopic scales [1]. It has found many applications such as in materials science and biomedical imaging [2]. Due to the high coherence and brilliance of synchrotron sources, resolutions down to tens of nanometers can be routinely achieved [3]. Several phase-sensitive imaging techniques have been developed [4], among them, propagation-based imaging [5] requires no equipment other than the source, object and detector. The absorption and phase shift induced by a sample can be retrieved from one or several such images, which is a nonlinear and ill-posed inverse problem.

Direct inversion formulas can be obtained by linearization of the problem [6–8]. Such analytical methods can work well under certain imaging conditions, but impose restrictions on the imaging setup [6] or the object [7, 8]. To overcome these limitations, for example in high resolution imaging [9], iterative algorithms that do not rely on linearization of the problem can be used. Among them are the methods of alternating projections on constraints in the object and the Fourier domain, e.g. the Hybrid Input-Output (HIO) algorithm [10]. Another class of algorithm is the so-called variational approaches that consist in minimizing a criterion defined by a data fidelity term and regularization terms. These algorithms permit a flexible inclusion of prior information, such as sparsity or Tikhonov regularization [11]. More recently, an iteratively regularised Gauss-Newton (IRGN) method was proposed, exploiting the regularity of the solution to obtain better robustness to high-frequency errors by regularizing with Sobolev norms [12]. Such approaches take into account the non-linearity of the problem, and yield better reconstructions with less artifacts than linearized methods. Unfortunately, the computation time is high and the choice of appropriate regularization remains a difficult problem.

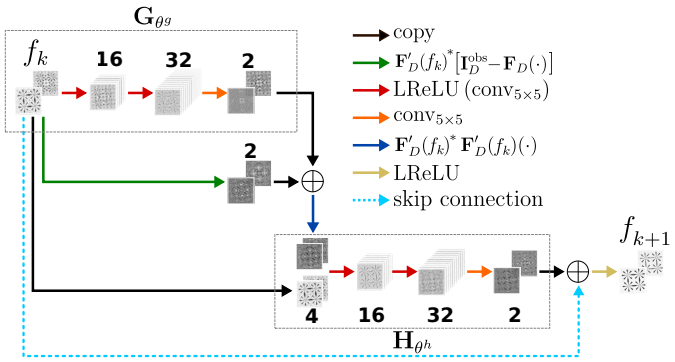

Fig. 1. Architecture of the network Γ_θ, representing one iteration of the Deep Gauss-Newton method.

The development of deep learning methods in recent years has led to many advances in image and signal processing [13]. Specifically, deep neural networks (DNNs) have been used to solve a wide variety of inverse problems [14]. Despite this progress, the black-box nature of DNNs, i.e., their lack of interpretability, is one of the primary obstacles for their use. Several approaches to exploiting DNNs for the solution of inverse problems have been proposed. DNNs can be used to reconstruct the unknown image directly from an available measurement, or some parts of an algorithm can be replaced by DNNs. Algorithm unrolling is an emerging technique based on the incorporation of convolutional neural networks (CNNs) into an iterative optimization scheme in order to give the DNN a specific role in the reconstruction [15]. Such approaches have found many applications [16] and unrolling has been applied to several optimization approaches: gradient descent algorithms [17], primal-dual schemes [18], and Alternating Direction Method of Multipliers [19]. For phase retrieval, several architectures have been proposed, including MS-D Net [20] and PhaseGAN [21].This kind of network is trained to approximate the inverse operator and often require large training sets as well as long training time. Other methods have incorporated neural networks into iterative schemes, but they are either computationally demanding [22, 23] or rely on a linearization of the forward model [24].

Table 1. Parameters for the Neural Networks

	MS-D Net	DGN
Loss function	MSE	MSE
Training epochs	100	100
Learning rate	10^{-3}	5×10^{-4}
Batch size	10	10
Activation function	ReLU	LReLU
Training time	35h	21h
Number of parameters	46×10^3	31×10^3

Here, we present a new learned iterative scheme, the Deep Gauss-Newton (DGN) algorithm, which is obtained by unrolling a Gauss-Newton iteration. The proposed method combines CNNs and knowledge of the imaging physics given by the forward operator and its Fréchet derivative. The rationale behind this choice is to take a well-known algorithm that is known to converge quickly and enhance it with machine learning by unrolling. We expect this scheme to inherit or improve the convergence properties of the Gauss-Newton method. Another advantage of this approach is that no regularization has to be chosen, instead it is adaptively learned from the data during training. We demonstrate the capability of the method to retrieve phase and attenuation from a single phase contrast image on simulated data as well as experimental data.

Under the projection approximation, the interaction of a coherent and parallel X-ray beam with matter can be described by a transmittance function [6]. In the experimental conditions considered here, the intensity at a distance D from the sample is well described by the Fresnel formalism, and is given by the convolution of the transmittance and the Fresnel propagator P_D:

$$\mathbf{F}_D(B,\varphi) = |\mathrm{e}^{-B+i\varphi} * P_D|^2 \quad \text{with} \quad P_D(\cdot) = \frac{1}{i\lambda D}\exp^{i\frac{\pi}{\lambda D}|\cdot|^2}. \quad \textbf{(1)}$$

The operator $\mathbf{F}_D$ describes the nonlinear relationship between the absorption B and phase shift φ induced by the object and the intensity of the diffraction pattern. In the following, we will note $f = (B, \varphi)$ the couple we aim to retrieve.

The inverse of $\mathbf{F}_D$ can be approximated using variational methods such as IRGN, corresponding to Tikhonov regularization of the Newton steps:

$$f_{k+1} = \underset{f}{\operatorname{argmin}} \left\{ \left|\left| \mathbf{F}_D(f_k) + \mathbf{F}'_D(f_k)\,(f - f_k) - \mathbf{I}_D^{\text{obs}} \right|\right|_2^2 + \alpha_k \,||f||_2^2 \right\} \quad \textbf{(2)}$$

where $\mathbf{F}'_D(f_k)$ is the Fréchet dérivative [25] of $\mathbf{F}_D$ at the point f_k, $\alpha_k > 0$ is a regularization parameter at iteration k, and $\mathbf{I}_D^{\text{obs}}$ is a noisy measured intensity. Eq. (2) has a unique solution

$$f_{k+1} = f_k + \left[\mathbf{F}'_D(f_k)^*\mathbf{F}'_D(f_k) + \alpha_k \mathrm{Id}\right]^{-1} \left\{ \mathbf{F}'_D(f_k)^* \left[\mathbf{I}_D^{\text{obs}} - \mathbf{F}_D(f_k)\right] - \alpha_k f_k \right\} \quad \textbf{(3)}$$

where $\mathbf{F}'_D(f_k)^*$ is the adjoint of the linear map $\mathbf{F}'_D(f_k)$ and Id is the identity. Usually, a step size for the update of f_k is introduced as regularization. Here, we propose instead to learn a regularization by replacing the Tikhonov term $\alpha_k f_k$ with a CNN $\mathbf{G}_{\theta_k^g}$ with parameters θ_k^g, and to approximate the inverse of the operator $\left[\mathbf{F}'_D(f_k)^*\mathbf{F}'_D(f_k) + \alpha_k \mathrm{Id}\right]$ with another CNN $\mathbf{H}_{\theta_k^h}$ with parameters θ_k^h, based on the current iterate f_k and on the approximate Hessian $\mathbf{F}'_D(f_k)^*\mathbf{F}'_D(f_k)$. The network H then replaces the classical approximation of the inverse of the Hessian used in the traditional Gauss-Newton scheme by a potentially better and faster learned approximation. If the algorithm is stopped after N iterations, we get $f_N = \left(\Gamma_{\theta_N} \circ \cdots \circ \Gamma_{\theta_1}\right)(f_0)$, where f_0 is the initial guess, $\theta_k = \left(\theta_k^g, \theta_k^h\right)$ and

$$\Gamma_{\theta_k}(f_k) = f_k + \mathbf{H}_{\theta_k^h}\left[f_k, \mathbf{F}'_D(f_k)^*\mathbf{F}'_D(f_k)\left\{\mathbf{F}'_D(f_k)^*\left[\mathbf{I}_D^{\text{obs}} - \mathbf{F}_D(f_k)\right] + \mathbf{G}_{\theta_k^g}(f_k)\right\}\right] \quad \textbf{(4)}$$

Unrolling this scheme, we can consider $\Lambda_\Theta = \Gamma_{\theta_N} \circ \cdots \circ \Gamma_{\theta_1}$ as a deep neural network representing N iterations with $\Theta = (\theta_1, \ldots, \theta_N)$ its parameters. Recent work on unrolling schemes has shown that using the same transformation at each iteration so that $\theta_k = \theta$ for $k \in \{1, \ldots, N\}$, yields good results [26]. Λ_Θ can then be seen as a recurrent neural network. The architecture of the network Γ_θ used for each iteration is shown in Fig. 1.

The network $\mathbf{G}_{\theta^g}$ takes the current iterate f_k as input, spreads it to 16 and then 32 channels by a convolutional layer with kernel size 5×5 using a leaky rectified linear unit (LReLU) as non-linearity, defined as $\mathrm{LReLU}_\alpha(x) = \max(x, \alpha x)$, $\alpha > 0$. The output of the network $\mathbf{G}_{\theta^g}$ is added to $\mathbf{F}'_D(f_k)^*\left[\mathbf{I}_D^{\text{obs}} - \mathbf{F}_D(f_k)\right]$, stacked with the current iterate and then fed to the network $\mathbf{H}_{\theta^h}$ which consists of the same set of operations as $\mathbf{G}_{\theta^g}$ (except it has four input channels instead of two). Finally, the output is added to the current iterate and projected onto positive numbers by a LReLU. The architectures of the networks are kept simple for several reasons. A shallow network added to the iterative update saves computational time and the memory required, while giving good reconstruction results.

To evaluate the algorithms, we created a database of 3D objects consisting of random combinations of one to ten ellipsoid and paraboloid shapes of three different materials (gold, palladium and zinc). We then computed 2D tomographic projections of the real and imaginary parts of the refractive index, corresponding to the phase shift and absorption. The X-ray energy was set to 13 keV for a wavelength of $\lambda = 0.095$ nm, and the propagation distance was set to $D = 20$ mm to yield a pixel size of 12 nm corresponding to the experimental conditions below. The phase contrast images were generated from the projection images according to the forward model Eq. (1). The propagation was calculated using 4 times oversampling with an image size of 2048×2048, yielding a final image size of 512×512 (Fig. 2). Gaussian noise with different amplitudes was added to the images to yield a peak to peak signal to noise ratio (PPSNR) between 10 and 100 dB. 12 000 pairs of one input image (simulated intensity) and two output images (absorption and phase) were generated. From this dataset, 10 000 pairs were used for training, 1 000 for validation during the training, and 1 000 for evaluation. We compare the unrolling framework to the standard IRGN method [12] as well as to the Mixed-Scale Dense Network (MS-D Net), a direct reconstruction method that does not include prior knowledge on the imaging physics [20].

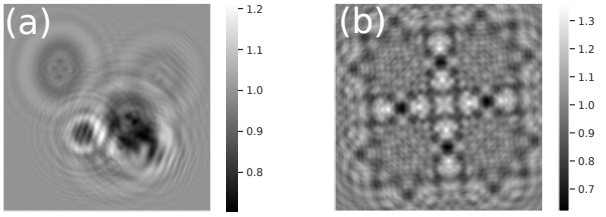

Fig. 2. **(a)** Simulated and **(b)** experimental intensity.

Table 2. Results (mean and standard deviation) on 1 000 simulated images.

Method	NMSE (%)		FRCM (%)		Resolution (nm)		Time (s)
	Absorption	Phase	Absorption	Phase	Absorption	Phase	
IRGN	85.5 (40.7)	39.3 (15.0)	71.2 (9.95)	68.1 (5.45)	238 (136)	154 (43)	116
MS-D Net	13.6 (12.8)	10.6 (10.8)	48.8 (13.8)	47.8 (13.3)	102 (77.4)	98.5 (135)	**2.60**
DGN	**12.1 (13.5)**	**4.61 (6.20)**	**35.7 (15.7)**	**23.0 (16.6)**	**72.2 (55.2)**	**62.3 (37.0)**	5.88

In the IRGN, the positive-definite linear operator in Eq. (3) can be inverted efficiently by a conjugate gradient (CG) method [12]. We used 100 Newton steps and 10 iterations for the CG. For the DGN method, we used $N = 10$ iterations, which means that the derivative $\mathbf{F}'_D(f_k)$ and its adjoint $\mathbf{F}'_D(f_k)^*$ are evaluated 10 times. The number of iterations was chosen empirically so that the NMSE stagnates. As opposed to [17], where several networks are trained sequentially, i.e. iteration by iteration, here, given a training set $\{y^i, f^i\}$ where y_i denotes the intensity $\mathbf{F}_D(f^i)$ corrupted by noise, we use one network $\mathbf{\Gamma}_\theta$ applied N times in a recurrent fashion to obtain the DNN $\mathbf{\Lambda}_\theta$, which is trained to perform end-to-end reconstruction. The DGN was trained using 100 epochs with a batch size of 10, the ADAM optimizer, an initial learning rate of 5×10^{-4} and a cosine annealing learning rate schedule [27]. The LReLU activation function parameter was set to the default $\alpha = 0.3$. Warm-up initialization decreased training time but did not yield better final results. Therefore, for simplicity, zero initialization, $f_0 = (0, 0)$, was used throughout. For the MS-D Net, we used the same settings as in [20]. Using LReLU in the MS-D Net did not improve the reconstructions. The hyperparameters for the networks are summarized in Tab. 1 and were optimized using grid search. Both DGN and MS-D Net were trained on the same training set.

To quantify reconstruction quality on simulated data, we used the normalized mean square error (NMSE). The frequency domain correlation between ground truth and recovered image was evaluated using the Fourier Ring Correlation (FRC). From this the Fourier Ring Correlation Metric (FRCM) [21] and the resolution were calculated. FRCM is the mean square difference between the FRC and unity over all spatial frequencies. A small FRCM implies a higher similarity in the Fourier domain. To assess the capability to retrieve high frequencies and reconstruct thin details, a metric to estimate the resolution was computed by the 2σ criterion [28]. The mean and standard deviation of the metrics were calculated on the test dataset. The average computation time for one reconstruction was measured to compare the execution time. The results are summarized in Tab. 2. Phase and absorption reconstructions from one simulated image pair are shown in Fig. 3. Both deep learning methods performed better than the IRGN method, which tended to leave artifacts in the absorption and yield a blurred phase. On average, the deep learning approaches performed better than the IRGN, both in terms of reconstructed values and resolution. In terms of NMSE, the deep learning approaches gave similar results for the absorption, but the DGN performed better than the MS-D Net for phase recovery. Moreover, the DGN yielded better resolution as well as better correlation in the frequency domain. As expected, the MS-D Net was fastest, since it only requires one application of a neural network. The DGN is efficient despite the need to compute the derivative of the foward operator as well as its adjoint several times and is 20 times faster than its standard iterative counterpart. For comparison, an unrolled simple gradient descent scheme [17] yielded less convincing results despite using more parameters (41×10^3): for the absorption 13.2 (17.3) % NMSE, 37.6 (15.2) % FRCM 82.2 (116) nm resolution and for the phase 4.74 (6.99) % NMSE, 23.9 (16.7) % FRCM, 69.3 (62.6) nm resolution, confirming our choice of network structure.

The proposed approach was applied on data acquired at beamline NanoMAX at the MAX IV synchrotron (Lund, Sweden) [29]. The diffraction pattern in Fig. 2 was obtained using the same image conditions as above. The sample consists of a stack of palladium, zinc, palladium, gold layers with thicknesses of 21, 10, 11, 163 nm, respectively, deposited on a 1 mm-thick silicon nitride substrate, resulting in expected values for absorption and phase of 0.0483 and 0.217 respectively. To evaluate the reconstructions quantitatively, we used the normalized error (NE) and relative standard deviation (RSD), calculated as $\text{NE} = \frac{l_t - l_m}{l_t}$ and $\text{RSD} = \frac{s_m}{l_m}$, where l_t is the expected value, l_m the measured mean value and s_m the standard deviation in the corresponding material. The shape of the object was estimated from a phase reconstruction using an iterative method, and chosen so that the calculation of l_m and s_m was done in homogeneous parts to avoid the influence of blur at the edges. The resolution was mea-

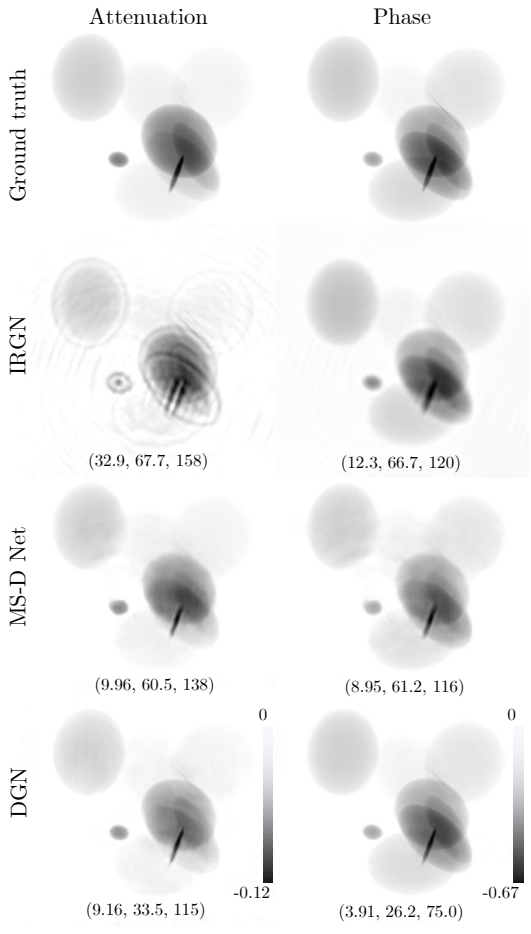

Fig. 3. Reconstructions from simulated data. Reconstruction quality is given as (NMSE (%), FRCM (%), resolution (nm)).

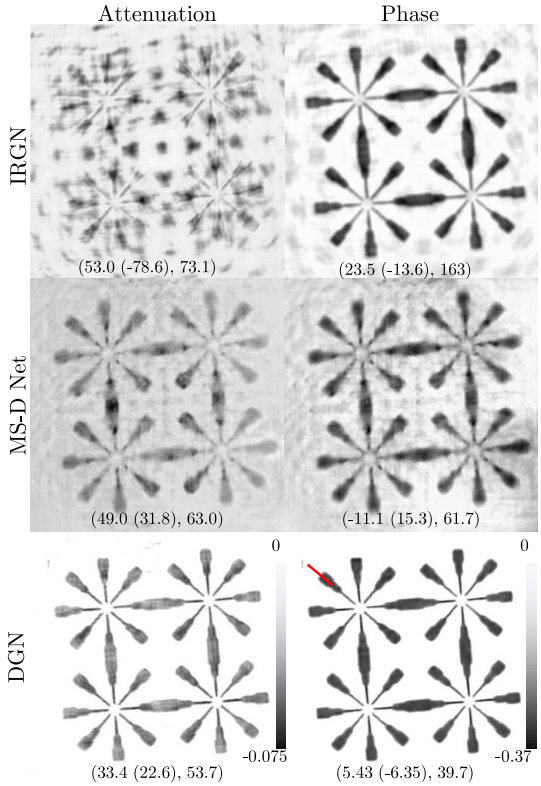

Fig. 4. Reconstructions for experimental data. The profiles along the red line were measured to estimate the resolution. Values correspond to (NE (%) (RSD (%)), resolution (nm)).

sured by fitting an error function to a line profile across an object edge and calculating the corresponding Gaussian full width at half maximum based on the error function fitting parameters.

The reconstructions obtained in Fig. 4 show that the DGN yields very high quality reconstructions of the object with almost no remaining visible artifacts. Although the MS-D Net performs well on simulations, it seems to not generalize as well as the DGN to the experimental data given the chosen training strategy. This could be due to the DGN explicitly taking into account the physics model, learning the noise statistics from the data yielding an adapted regularization, while leveraging the convergence properties of the optimization method [16]. On the other hand, the MS-D Net was trained to reconstruct directly from the measurements, without knowledge of the physics model.

A limitation of the proposed algorithm is that the forward model has to be fully known, e.g. the propagation distance has to be precisely measured. In future work we will study possibilites to correct errors in the forward model. We will also investigate out-of-distribution generalization error, e.g. with respect to different noise properties. The developed DGN algorithm allows to efficiently retrieve both the phase and absorption from a single phase contrast image. By exploiting recent developments in deep learning and integrating CNNs into a regularized Gauss-Newton scheme, with the DGN we overcome the limitations of classical iterative approaches while leveraging the power of neural networks. Since no regularization term needs to be chosen, the DGN network is trained to learn an optimal one for the absorption and the phase respectively, which improves the quality of the reconstructions. Taking into account the knowledge of the forward model in a simple network enhances the reconstructions and allows a better generalization on real data. Compared to the standard IRGN, the DGN method both substantially improved the reconstruction and reduced the calculation time.

Acknowledgments. Pablo Villanueva-Perez (Lund University, Lund, Sweden) and Sebastian Kalbfleich (MAX IV Laboratory, Lund, Sweden) for the experimental data, and Jesper Wallentin and Lert Chayanun (NanoLund and Lund University, Lund, Sweden) for the sample.

Disclosures. The authors declare no conflicts of interest.

Data Availability Statement. Data underlying the simulated results presented in this paper can be generated using *TomoPhantom* software [30]. The data underlying the experimental results presented are available through the *PyPhase* package [31].

REFERENCES

1. M. Langer and F. Peyrin, Osteoporos. Int. **27**, 441 (2016).
2. M. Langer, *In-Line X-Ray Phase Tomography of Bone and Biomaterials for Regenerative Medicine* (Springer, Cham, 2018).
3. J.-P. Suuronen, B. Hesse, M. Langer, M. Bohner, and J. Villanova, J. Synchrotron Radiat. **29**, 843 (2022).
4. S. Tao, C. He, C. Kuang, and X. Liu, Appl. Sci. **11**, 2971 (2021).
5. A. Snigirev, I. Snigireva, V. Kohn, S. Kuznetsov, and I. Schelokov, Rev. Sci. Instruments **66**, 5486 (1995).
6. D. M. Paganin, *Coherent X-ray optics* (Oxford University Press, 2006).
7. P. Cloetens, W. Ludwig, J. Baruchel, D. Van Dyck, J. Van Landuyt, J. P. Guigay, and M. Schlenker, Appl. Phys. Lett. **75**, 2912 (1999).
8. J. P. Guigay, M. Langer, R. Boistel, and P. Cloetens, Opt. Lett. **32**, 1617 (2007).
9. R. Mokso, P. Cloetens, E. Maire, J.-Y. Buffiere, and W. Ludwig, Applied Physics Letters p. 1 (2007).
10. J. R. Fienup, Appl. Opt. **21 (15)**, 2758 (1982).
11. V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, Appl. Opt. **52**, 3977 (2013).
12. S. Maretzke, M. Bartels, M. Krenkel, T. Salditt, and T. Hohage, Opt. Express **24**, 6490 (2016).
13. Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
14. G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, IEEE J. on Sel. Areas Inf. Theory **1**, 39 (2020).
15. S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb, Acta Numer. **28**, 1–174 (2019).
16. V. Monga, Y. Li, and Y. C. Eldar, IEEE Signal Proc. Mag. **38**, 18 (2021).
17. A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, and S. Arridge, IEEE Trans. on Med. Imaging **37**, 1382 (2018).
18. J. Adler and O. Öktem, IEEE Trans. on Med. Imaging **37**, 1322 (2018).
19. Y. Yang, J. Sun, H. Li, and Z. Xu, IEEE Transactions on Pattern Analysis Mach. Intell. **42**, 521 (2020).
20. K. Mom, B. Sixou, and M. Langer, Appl. Opt. **61**, 2497 (2022).
21. Y. Zhang, M. A. Noack, P. Vagovic, K. Fezzaa, F. Garcia-Moreno, T. Ritschel, and P. Villanueva-Perez, Opt. Express **29**, 19593 (2021).
22. C. A. Metzler, P. Schniter, A. Veeraraghavan, and R. G. Baraniuk, CoRR **abs/1803.00212** (2018).
23. C. Işil, F. S. Oktem, and A. Koç, Appl. Opt. **58**, 5422 (2019).
24. C. Bai, M. Zhou, J. Min, S. Dand, X. Yu, P. Zhang, T. Peng, and B. Yao, Opt. Lett. **44 (21)**, 5141 (2019).
25. V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, Opt. Express **19**, 22809 (2011).
26. R. Dabre and A. Fujita, AAAI Conf. on Artif. Intell. **33**, 6292 (2019).
27. I. Loshchilov and F. Hutter, "Sgdr: Stochastic gradient descent with warm restarts," in *ICLR 2017,* (2016).
28. N. Banterle, K. H. Bui, E. A. Lemke, and M. Beck, J. Struct. Biol. **183**, 363 (2013).
29. S. Kalbfleisch, M. Kahnt, K. Buakor, M. Langer, T. Dreier, H. Dierks, P. Stjärneblad, E. Larsson, K. Gordeyeva, L. Chayanun, D. Söderberg, J. Wallentin, M. Bech, and P. Villanueva-Perez, J. Synch. Radiat. **29** (2022).
30. D. Kazantsev, V. Pickalov, S. Nagella, P. Edoardo, and P. J. Withers, SoftwareX **7**, 150 (2018).
31. M. Langer, Y. Zhang, D. Figueirinhas, J.-B. Forien, K. Mom, C. Mouton, R. Mokso, and P. Villanueva-Perez, J. Synch. Radiat. **28**, 1261 (2021).

FULL REFERENCES

1. M. Langer and F. Peyrin, "3D X-ray ultra-microscopy of bone tissue," Osteoporos. Int. **27**, 441–455 (2016).
2. M. Langer, *In-Line X-Ray Phase Tomography of Bone and Biomaterials for Regenerative Medicine* (Springer, Cham, 2018).
3. J.-P. Suuronen, B. Hesse, M. Langer, M. Bohner, and J. Villanova, "Evaluation of imaging setups for quantitative phase contrast nanoCT of mineralized biomaterials," J. Synchrotron Radiat. **29**, 843–852 (2022).
4. S. Tao, C. He, C. Kuang, and X. Liu, "Principles of different x-ray phase-contrast imaging: A review," Appl. Sci. **11**, 2971 (2021).
5. A. Snigirev, I. Snigireva, V. Kohn, S. Kuznetsov, and I. Schelokov, "On the possibilities of x-ray phase contrast microimaging by coherent high-energy synchrotron radiation," Rev. Sci. Instruments **66**, 5486–5492 (1995).
6. D. M. Paganin, *Coherent X-ray optics* (Oxford University Press, 2006).
7. P. Cloetens, W. Ludwig, J. Baruchel, D. Van Dyck, J. Van Landuyt, J. P. Guigay, and M. Schlenker, "Holotomography: Quantitative phase tomography with micrometer resolution using hard synchrotron radiation x rays," Appl. Phys. Lett. **75**, 2912–2914 (1999).
8. J. P. Guigay, M. Langer, R. Boistel, and P. Cloetens, "Mixed transfer function and transport of intensity approach for phase retrieval in the fresnel region," Opt. Lett. **32**, 1617–1619 (2007).
9. R. Mokso, P. Cloetens, E. Maire, J.-Y. Buffiere, and W. Ludwig, "Nanoscale zoom tomography with hard x-rays using kirkpatrick-baez optics," Applied Physics Letters p. 1 (2007).
10. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21 (15)**, 2758–2769 (1982).
11. V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, "Nonlinear approaches for the single-distance phase retrieval problem involving regularizations with sparsity constraints," Appl. Opt. **52**, 3977–3986 (2013).
12. S. Maretzke, M. Bartels, M. Krenkel, T. Salditt, and T. Hohage, "Regularized newton methods for x-ray phase contrast and general imaging problems," Opt. Express **24**, 6490–6506 (2016).
13. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–44 (2015).
14. G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," IEEE J. on Sel. Areas Inf. Theory **1**, 39–56 (2020).
15. S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb, "Solving inverse problems using data-driven models," Acta Numer. **28**, 1–174 (2019).
16. V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," IEEE Signal Proc. Mag. **38**, 18–44 (2021).
17. A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, and S. Arridge, "Model-based learning for accelerated, limited-view 3-d photoacoustic tomography," IEEE Trans. on Med. Imaging **37**, 1382–1393 (2018).
18. J. Adler and O. Öktem, "Learned primal-dual reconstruction," IEEE Trans. on Med. Imaging **37**, 1322–1332 (2018).
19. Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-csnet: A deep learning approach for image compressive sensing," IEEE Transactions on Pattern Analysis Mach. Intell. **42**, 521–538 (2020).
20. K. Mom, B. Sixou, and M. Langer, "Mixed scale dense convolutional networks for x-ray phase contrast imaging," Appl. Opt. **61**, 2497–2505 (2022).
21. Y. Zhang, M. A. Noack, P. Vagovic, K. Fezzaa, F. Garcia-Moreno, T. Ritschel, and P. Villanueva-Perez, "PhaseGAN: a deep-learning phase-retrieval approach for unpaired datasets," Opt. Express **29**, 19593–19604 (2021).
22. C. A. Metzler, P. Schniter, A. Veeraraghavan, and R. G. Baraniuk, "prdeep: Robust phase retrieval with flexible deep neural networks," CoRR **abs/1803.00212** (2018).
23. C. Işil, F. S. Oktem, and A. Koç, "Deep iterative reconstruction for phase retrieval," Appl. Opt. **58**, 5422–5431 (2019).
24. C. Bai, M. Zhou, J. Min, S. Dand, X. Yu, P. Zhang, T. Peng, and B. Yao, "Robust contrast-transfer-function phase retrieval via flexible deep learning networks," Opt. Lett. **44 (21)**, 5141–5144 (2019).
25. V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, "Non-linear iterative phase retrieval based on frechet derivative," Opt. Express **19**, 22809–22819 (2011).
26. R. Dabre and A. Fujita, "Recurrent stacking of layers for compact neural machine translation models," AAAI Conf. on Artif. Intell. **33**, 6292–6299 (2019).
27. I. Loshchilov and F. Hutter, "Sgdr: Stochastic gradient descent with warm restarts," in *ICLR 2017,* (2016).
28. N. Banterle, K. H. Bui, E. A. Lemke, and M. Beck, "Fourier ring correlation as a resolution criterion for super-resolution microscopy," J. Struct. Biol. **183**, 363–367 (2013).
29. S. Kalbfleisch, M. Kahnt, K. Buakor, M. Langer, T. Dreier, H. Dierks, P. Stjärneblad, E. Larsson, K. Gordeyeva, L. Chayanun, D. Söderberg, J. Wallentin, M. Bech, and P. Villanueva-Perez, "X-ray in-line holography and holotomography at the nanomax beamline," J. Synch. Radiat. **29** (2022).
30. D. Kazantsev, V. Pickalov, S. Nagella, P. Edoardo, and P. J. Withers, "TomoPhantom, a software package to generate 2D–4D analytical phantoms for CT image reconstruction algorithm benchmarks," SoftwareX **7**, 150–155 (2018).
31. M. Langer, Y. Zhang, D. Figueirinhas, J.-B. Forien, K. Mom, C. Mouton, R. Mokso, and P. Villanueva-Perez, "*PyPhase* – a Python package for X-ray phase imaging," J. Synch. Radiat. **28**, 1261–1266 (2021).