
Learning Regularization Functionals for Inverse Problems: A Comparative Study

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 In recent years, a variety of learned regularization frameworks for solving inverse
2 problems in imaging have emerged. These offer flexible modeling together with
3 mathematical insights. The proposed methods differ in their architectural design
4 and training strategies, making direct comparison challenging due to non-modular
5 implementations. We address this gap by collecting and unifying the available
6 code into a common framework. This unified view allows us to systematically
7 benchmark the approaches and highlight their strengths and limitations, providing
8 valuable insights into their future potential.

9 1 Introduction

10 Inverse problems are ubiquitous in imaging sciences, for example for describing the image acquisition
11 process of magnetic resonance imaging (MRI) or computed tomography (CT). Mathematically, the
12 reconstruction is modeled as a linear inverse problem. More precisely, we want to reconstruct an
13 (unknown) image $\mathbf{x} \in \mathbb{R}^d$ from an observation $\mathbf{y} \in \mathbb{R}^m$ determined by the linear relation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

14 where $\mathbf{H} \in \mathbb{R}^{m \times d}$ encodes the underlying data acquisition process and the noise $\mathbf{n} \in \mathbb{R}^m$ accounts
15 for imperfections in this description. As \mathbf{H} is often ill-conditioned or non-invertible, the inverse
16 problem (1) is ill-posed in the sense of Hadamard and reconstructing \mathbf{x} from \mathbf{y} is challenging. A
17 classical method to address ill-posedness is variational regularization [39], for which the unknown \mathbf{x}
18 is approximated by

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \min_{\mathbf{x}} \{D(\mathbf{H}\mathbf{x}, \mathbf{y}) + \alpha R(\mathbf{x})\}. \quad (2)$$

19 In (2), the data fidelity $D: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ ensures data consistency, the regularizer $R: \mathbb{R}^d \rightarrow \mathbb{R}$
20 promotes desired properties of \mathbf{x} , and the regularization parameter $\alpha > 0$ balances the two. The
21 literature on variational regularization methods and their mathematical analysis is vast, see [8, 22, 39]
22 and the references therein. In particular, the variational approach (2) leads to several desirable
23 properties like universality, data consistency, stability and interpretability. Over the past years,
24 deep-learning-based approaches have become the state-of-the-art for solving inverse problems [7].
25 However, several concerns regarding their trustworthiness remain, see, e.g., [4, 16]. In contrast,
26 hand-crafted regularizers R such as the total variation (TV) [37] are theoretically founded but cannot
27 achieve the same reconstruction quality as data-driven approaches. We focus on the blend of these
28 approaches, namely the learning of R from data.

29 **Contributions** We provide an overview of architectures and training methods for learned regular-
30 izers and provide a systematic benchmark in a unified setting. We provide implementations to all
31 examples and highlight the individual strengths and limitations.

2 Architectures and Training Methods

Architectures A pioneering learnable regularizer R is the Field of Experts (FoE) [36], which is the sum of 1D potentials composed with convolutional filters. Recently, it was proposed to learn the FoE using linear splines, leading to the Convex Ridge Regularizer (CRR) [17] and Weakly Convex Ridge Regularizer (WCRR) [18]. Another convex architecture is the Input Convex Neural Network (ICNN) [7] and its descendant, namely the input weakly-convex neural network [40]. Following the idea of structured nonconvexity, these were extended to Input Difference of Convex Neural Networks (IDCNNs) [44]. Examples of more complex multiscale CNN regularizers are the Total Deep Variation (TDV) [25] and the Least Squares Residual (LSR) [46]. Exploiting the self-similarity of natural and medical images, the expected patch log likelihood (EPLL) [45] and patch normalizing flow regularizers (PatchNR) [3] define a regularizer by representing the patch distribution by a mixture model or generative model. Finally, some Plug-and-Play frameworks provably define a regularization term R , see [14, 21]. As an example, we include Learned Proximal Networks (LPNs) [14]. Other architectures, which are not included in our comparison, are energy-based generative priors [43], dictionary learning [11, 32, 42] and priors based on generative models [2, 10, 13, 20].

Training Methods First, we consider bilevel learning (BL), which adapts the θ such that the reconstruction (2) minimizes some loss. This idea started with learning only the regularization parameter α in (2) [12, 19, 26], and has been gradually lifted to learning regularizers. Here, we rely on Jacobian free backpropagation [9, 15] to compute the required gradients. Alternatively, the gradients can be computed by implicit differentiation [23, 30], which also allows to choose the involved step sizes and accuracies adaptively to ensure convergence [38]. A second paradigm is based on distinguishing desirable and undesirable images a priori, without actually solving (1). Prominent examples include (local) adversarial regularization (AR) [31, 34, 35] and network Tikhonov regularizers (NETT) [29]. During training, these approaches are not linked to the variational problem (2), and require the selection of a suitable regularization parameter α for the inverse problem at hand. Finally, from a Bayesian viewpoint, regularizers can be characterized by the log density of the training dataset. Using this identity, R can be trained via maximum likelihood losses, which we use for the patch-based architectures EPLL and PatchNR. Alternatively, we can use a score matching objective. While this is computationally very efficient, we found that the results are not competitive towards BL and AR. Instead, we use it as a pretraining method to initialize θ for BL.

3 Comparative Study

We perform an experimental comparison of the architectures and training methods for CT reconstruction. To solve (2) efficiently, we use the nonmonotonic Accelerated Proximal Gradient algorithm [28, Supplementary material]. Our code is based on the DeepInverse library [41] and available online¹.

Forward Operator and Noise Level We consider a sparse-view CT setting, where \mathbf{H} is given by the discretized X-ray transform with 60 equispaced angles and a parallel beam geometry. We use the DeepInverse implementation. To keep the setup simple, we consider Gaussian noise with $\sigma = 0.7$ instead of more realistic Poisson noise.

Datasets and Experiments We evaluate all models on the first test batch of the LoDoPaB-CT dataset [27] consisting out of 128 images of size 362×362 based on the LIDC/IDRI database [6]. For the training phase, we consider two different settings. In the *supervised CT* setting, we assume that we have access to training images from LoDoPaB-CT. To this end, we use 3522 images from its validation set for training. However, in practice, we rarely have training images from the same domain or scanner. Therefore, we also consider an *unsupervised CT* setting, where we train the regularizer R based on natural images from the BSDS500 dataset [5, 33]. In this case, we also train on a denoising task without actually using the forward operator \mathbf{H} . As an error measure, we report the peak signal-to-noise ratio.

Results The results for supervised CT are given in Table 1 and for unsupervised CT in Table 2. Additionally, Table 3 contains training methods that are specific to certain architectures. Moreover,

¹link anonymized

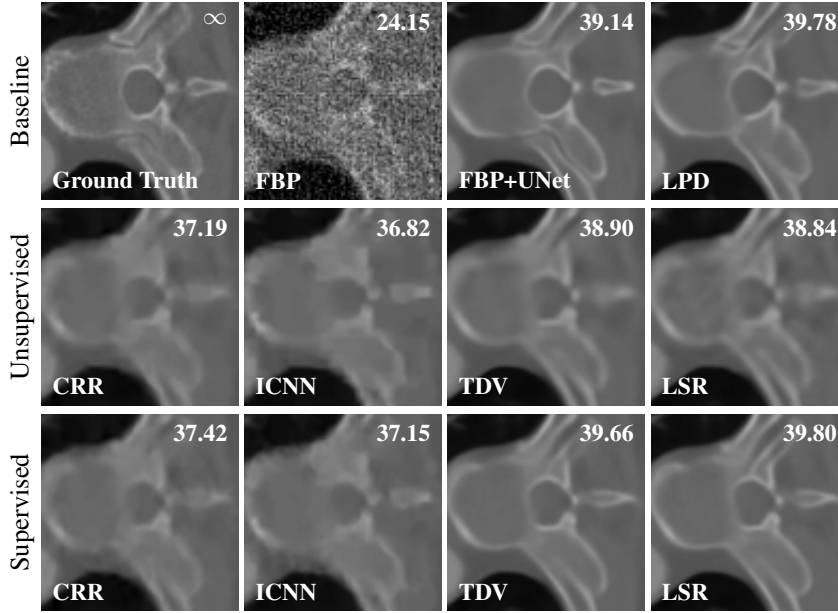


Figure 1: Reconstructions (crop) for the CT task for some baselines and BL with various architectures.

Table 1: Supervised CT experiment.

	CRR	ICNN	WCRR	IDCNN	CNN	TDV	LSR
BL	32.30	32.16	32.85	32.56	-	33.67	33.72
AR	32.23	31.98	32.48	31.93	32.29	32.33	-

Table 2: Unsupervised CT experiment.

	CRR	ICNN	WCRR	IDCNN	CNN	TDV	LSR
BL	32.17	31.99	32.65	32.45	32.69	33.23	33.11
AR	32.14	31.94	32.61	31.98	32.04	32.43	-

Table 3: Baselines and regularizers with architecture-specific training routines.

	FBP	TV	FBP+UNET	LPD	EPLL	PatchNR	NETT	LPN
unsupervised	19.98	30.99	n/a	n/a	31.94	32.17	30.64	31.29
supervised	n/a	n/a	33.03	33.71	32.55	32.63	32.01	32.08

Table 3 includes common baselines such as the filtered back projection (FBP), the TV reconstruction, a UNet-based postprocessing of the FBP (FBP+UNet) [24] and a learned primal dual reconstruction (LPD) [1]. A hyphen indicates that the corresponding setup was unstable or required a very long training time. Visual reconstruction results for BL and some baselines are provided in Figure 1.

4 Discussion, Conclusions and Limitations

From our experiments, we conclude that learned regularizers like the TDV and LSR can produce competitive results to LPD, which is state-of-the-art for CT. Also the (weakly) convex architectures like (W)CRR and ICNN significantly outperform classical approaches like TV while still providing theoretical guarantees. Interestingly, the CRR consistently yields a higher PSNR than the more general ICNN architecture. Considering the training methods, BL always outperformed AR in terms of the PSNR. On the other hand, we note that AR is much faster to train, even though we do not report systematic results in this direction. Finally, all regularizers perform well in the unsupervised setting, which is a clear advantage over end-to-end reconstruction networks. Nevertheless, there is a slight drop in PSNR and also the visual impression degrades. The reconstruction times are competitive with other iterative reconstruction methods such as the well established PnP approach.

While our results give a clear impression of the capabilities and limitations of the specific architectures and training methods, several aspects remain to be explored. This includes robustness towards errors in the forward model \mathbf{H} , the required size of the dataset, and uncertainty estimates. Also it is unclear if training R on several inverse problems can lead to a foundational model for inverse problems.

References

- [1] J. Adler and O. Öktem. Learned primal-dual reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1322–1332, 2018.
- [2] G. S. Alberti, J. Hertrich, M. Santacesaria, and S. Sciutto. Manifold learning by mixture models of VAEs for inverse problems. *Journal of Machine Learning Research*, 25(202):1–35, 2024.
- [3] F. Altekürger, A. Denker, P. Hagemann, J. Hertrich, P. Maass, and G. Steidl. PatchNR: Learning from very few images by patch normalizing flow regularization. *Inverse Problems*, 39(6):064006, 2023.
- [4] V. Antun, F. Renna, C. Poon, B. Adcock, and A. C. Hansen. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proceedings of the National Academy of Sciences*, 117(48):30088–30095, 2020.
- [5] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916, 2011.
- [6] S. G. I. Armato, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*, 38(2):915–931, 2011.
- [7] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019.
- [8] M. Benning and M. Burger. Modern Regularization Methods for Inverse Problems. *Acta Numerica*, 27:1–111, 2018.
- [9] J. Bolte, E. Pauwels, and S. Vaiter. One-step differentiation of iterative algorithms. In *Advances in Neural Information Processing Systems 36*. Curran Associates, Inc., 2023.
- [10] A. Bora, A. Jalal, E. Price, and A. G. Dimakis. Compressed sensing using generative models. In *Proceedings of the 34th International Conference on Machine Learning*, pages 537–546. PMLR, 2017.
- [11] A. M. Bruckstein, D. L. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, 2009.
- [12] J. C. De los Reyes and C.-B. Schönlieb. Image denoising: Learning the noise model via nonsmooth PDE-constrained optimization. *Inverse Problems and Imaging*, 7(4):1183–1214, 2013.
- [13] M. A. G. Duff, N. D. F. Campbell, and M. J. Ehrhardt. Regularising inverse problems with generative machine learning models. *Journal of Mathematical Imaging and Vision*, 66:37–56, 2024.
- [14] Z. Fang, S. Buchanan, and J. Sulam. What’s in a prior? Learned proximal networks for inverse problems. In *International Conference on Learning Representations*, 2024.
- [15] S. W. Fung, H. Heaton, Q. Li, D. McKenzie, S. Osher, and W. Yin. JFB: Jacobian-free backpropagation for implicit networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(6):6648–6656, 2022.
- [16] N. M. Gottschling, V. Antun, A. C. Hansen, and B. Adcock. The troublesome kernel: On hallucinations, no free lunches, and the accuracy-stability tradeoff in inverse problems. *SIAM Review*, 67(1):73–104, 2025.
- [17] A. Goujon, S. Neumayer, P. Bohra, S. Ducotterd, and M. Unser. A neural-network-based convex regularizer for inverse problems. *IEEE Transactions on Computational Imaging*, 9:781–795, 2023.

- [18] A. Goujon, S. Neumayer, and M. Unser. Learning weakly convex regularizers for convergent image-reconstruction algorithms. *SIAM Journal on Imaging Sciences*, 17(1):91–115, 2024.
- [19] E. Haber, L. Horesh, and L. Tenorio. Numerical methods for the design of large-scale nonlinear discrete ill-posed inverse problems. *Inverse Problems*, 26(2):025002, 2009.
- [20] A. Habring and M. Holler. A generative variational model for inverse problems in imaging. *SIAM Journal on Mathematics of Data Science*, 4(1):306–335, 2022.
- [21] S. Hurault, A. Leclaire, and N. Papadakis. Proximal denoiser for convergent Plug-and-Play optimization with nonconvex regularization. In *Proceedings of the 39th International Conference on Machine Learning*, pages 9483–9505. PMLR, 2022.
- [22] K. Ito and B. Jin. *Inverse Problems: Tikhonov Theory and Algorithms*. World Scientific, 2014.
- [23] K. Ji, J. Yang, and Y. Liang. Bilevel optimization: Convergence analysis and enhanced design. In *Proceedings of the 38th International Conference on Machine Learning*, pages 4882–4892. PMLR, 2021.
- [24] 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.
- [25] E. Kobler, A. Effland, K. Kunisch, and T. Pock. Total deep variation for linear inverse problems. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
- [26] K. Kunisch and T. Pock. A bilevel optimization approach for parameter learning in variational models. *SIAM Journal on Imaging Sciences*, 6(2):938–983, 2013.
- [27] J. Leuschner, M. Schmidt, D. O. Baguer, and P. Maass. LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction. *Scientific Data*, 8(1):109, 2021.
- [28] H. Li and Z. Lin. Accelerated proximal gradient methods for nonconvex programming. In *Advances in Neural Information Processing Systems 28*. Curran Associates, Inc., 2015.
- [29] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier. NETT: Solving inverse problems with deep neural networks. *Inverse Problems*, 36(6):065005, 2020.
- [30] R. Liao, Y. Xiong, E. Fetaya, L. Zhang, K. Yoon, X. Pitkow, R. Urtasun, and R. Zemel. Reviving and improving recurrent back-propagation. In *Proceedings of the 35th International Conference on Machine Learning*, pages 3082–3091. PMLR, 2018.
- [31] S. Lunz, O. Öktem, and C.-B. Schönlieb. Adversarial regularizers in inverse problems. In *Advances in Neural Information Processing Systems 31*. Curran Associates, Inc., 2018.
- [32] J. Mairal, J. Ponce, G. Sapiro, A. Zisserman, and F. Bach. Supervised dictionary learning. In *Advances in Neural Information Processing Systems 21*. Curran Associates, Inc., 2008.
- [33] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th International Conference on Computer Vision*, pages 416–423, 2001.
- [34] S. Mukherjee, S. Dittmer, Z. Shumaylov, S. Lunz, O. Öktem, and C.-B. Schönlieb. Learned convex regularizers for inverse problems. *arXiv:2008.02839*, 2021.
- [35] J. Prost, A. Houdard, A. Almansa, and N. Papadakis. Learning local regularization for variational image restoration. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 358–370. Springer, 2021.
- [36] S. Roth and M. J. Black. Fields of experts. *International Journal of Computer Vision*, 82(2):205–229, 2009.
- [37] 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.

- 191 [38] M. S. Salehi, S. Mukherjee, L. Roberts, and M. J. Ehrhardt. An adaptively inexact first-order
192 method for bilevel optimization with application to hyperparameter learning. *SIAM Journal on*
193 *Mathematics of Data Science*, 7(3):906–936, 2025.
- 194 [39] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in*
195 *Imaging*. Springer, 2009.
- 196 [40] Z. Shumaylov, J. Budd, S. Mukherjee, and C.-B. Schönlieb. Weakly convex regularisers for
197 inverse problems: Convergence of critical points and primal-dual optimisation. In *Proceedings*
198 *of the 41st International Conference on Machine Learning*. PMLR, 2024.
- 199 [41] J. Tachella, M. Terris, S. Hurault, A. Wang, D. Chen, M.-H. Nguyen, M. Song, T. Davies,
200 L. Davy, J. Dong, P. Escande, J. Hertrich, Z. Hu, T. I. Liaudat, N. Laurent, B. Levac, M. Mas-
201 sias, T. Moreau, T. Modrzyk, B. Monroy, S. Neumayer, J. Scanvic, F. Sarron, V. Sechaud,
202 G. Schramm, C. Tang, R. Vo, and P. Weiss. DeepInverse: A Python package for solving imaging
203 inverse problems with deep learning. *arXiv preprint arXiv:2505.20160*, 2025.
- 204 [42] I. Tošić and P. Frossard. Dictionary learning. *IEEE Signal Processing Magazine*, 28(2):27–38,
205 2011.
- 206 [43] M. Zach, E. Kobler, and T. Pock. Computed tomography reconstruction using generative
207 energy-based priors. In *OAGM Workshop*, pages 52–58, 2021.
- 208 [44] Y. Zhang and O. Leong. Learning difference-of-convex regularizers for inverse problems: A
209 flexible framework with theoretical guarantees. *arXiv preprint arXiv:2502.00240*, 2025.
- 210 [45] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image
211 restoration. In *IEEE International Conference on Computer Vision*, pages 479–486, 2011.
- 212 [46] Z. Zou, J. Liu, B. Wohlberg, and U. S. Kamilov. Deep equilibrium learning of explicit regu-
213 larization functionals for imaging inverse problems. *IEEE Open Journal of Signal Processing*,
214 4:390–398, 2023.