SUB-DIP: Optimization on a subspace with deep image prior regularization and application to superresolution

Alexander Sagel, Aline Roumy, Christine Guillemot

SUB-DIP: OPTIMIZATION ON A SUBSPACE WITH DEEP IMAGE PRIOR REGULARIZATION AND APPLICATION TO SUPERRESOLUTION

Alexander Sagel

Aline Roumy, Christine Guillemot

fortiss

Munich, Germany

Inria

Rennes, France

ABSTRACT

The Deep Image Prior has been recently introduced to solve inverse problems in image processing with no need for training data other than the image itself. However, the original training algorithm of the Deep Image Prior constrains the reconstructed image to be on a manifold described by a convolutional neural network. For some problems, this neglects prior knowledge and can render certain regularizers ineffective. This work proposes an alternative approach that relaxes this constraint and fully exploits all prior knowledge. We evaluate our algorithm on the problem of reconstructing a high-resolution image from a downsampled version and observe a significant improvement over the original Deep Image Prior algorithm.

Index Terms— Image reconstruction, Image restoration, Inverse problems, Neural networks

1. INTRODUCTION

Inverse problems appear in a variety of image processing applications, e.g., image denoising, inpainting or superresolution. A common strategy for dealing with ill-posed inverse problems consists in introducing some prior knowledge on the kind of typical images we try to restore, which helps restricting the class of admissible solutions.

In their simplest form, priors are handcrafted regularizers that promote certain properties assumed about the data at hand. A prominent example is *Total Variation* (TV) regularization [1] that favors piecewise smooth solutions. More sophisticated priors require some form of learning procedure applied to the input image in order to infer additional structural information. For instance, it is common to assume that localized patches of a natural image have a sparse representation w.r.t. a learned dictionary [2].

Nowadays, deep learning based approaches are considered state of the art at solving inverse image processing problems, including superresolution [3] and inpainting [4]. How-

ever, unlike the traditional "shallow" algorithms mentioned above, they entirely rely on very large data sets.

The *Deep Image Prior* (DIP) [5] has recently demonstrated that a deep convolutional neural network (CNN) can still provide superior performance on different kinds of inverse problems, even though it is trained exclusively on the input image itself, thus leveraging the advantages of traditional and deep learning algorithms.

In the original formulation, DIP searches for a CNN-generated image that fulfills best the constraint posed by the inverse problem. Thus, it models the set of possible solutions as the manifold described by the CNN and the constraints posed by the inverse problem as a real-valued energy function to be minimized. Our work proposes the opposite point of view. The feasible set is defined by the inverse problem, while DIP is merely a regularizing energy function to be minimized. We show that this inversion of perspective can improve DIP performance on superresolution problems.

2. DIP: BACKGROUND AND PRIOR WORK

Many inverse problems consist of reconstructing an image $\boldsymbol{x} \in \mathbb{R}^n$ from a measurement $\boldsymbol{y} \in \mathbb{R}^d$, $d < n$, i.e.

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\eta}, \tag{1}$$

where $\boldsymbol{A}$ represents a linear *degradation operator*. The term η corresponds to an optional Gaussian noise.

DIP approximates the solution of the underdetermined linear Eq. (1) by employing a CNN. Specifically, given a CNN

$$T_\theta : \mathbb{R}^{\tilde{n}} \to \mathbb{R}^n \tag{2}$$

parameterized by its trainable weights $\theta \in \mathbb{R}^p$, and a fixed input $\boldsymbol{z} \in \mathbb{R}^{\tilde{n}}$, DIP estimates the minimizer θ^* of the objective

$$f_{\mathrm{DIP}}(\theta) = \|\boldsymbol{A}T_\theta(\boldsymbol{z}) - \boldsymbol{y}\|^2, \tag{3}$$

and returns

$$\boldsymbol{x}^*_{\mathrm{DIP}} = T_{\theta^*}(\boldsymbol{z}), \text{ s.t. } \theta^* = \arg\min_{\theta \in \mathbb{R}^p} f_{\mathrm{DIP}}(\theta) \tag{4}$$

as the approximated solution of Eq. (1).

Even though the DIP model was introduced only recently, some important research focusing on the formulation of its

optimization has been already carried out. Several interpretations of its training objective in Eq. (3) have been discussed and theoretically analyzed in [6]. Of interest is the question if DIP can be further improved by adjusting the training objective. For instance, regularization of DIP has been discussed in [7] for TV and in [8] for learned regularization. Similarly, the work [9] proposes to combine DIP with the concept of regularization by denoising introduced in [10]. A peculiarity of this work is its optimization algorithm. Instead of optimizing the defined objective directly via some variation of gradient descent, as is common for deep neural networks, an alternating direction method of multipliers (ADMM) is chosen.

Like the aforementioned works, our work also discusses the DIP objective and proposes a way to improve it. However, aside from additional regularization terms, the discussed prior works stick to the original (or an equivalent) DIP formulation of the optimization objective, while the contribution of our work is to rephrase this objective.

3. A SUBSPACE INDUCED DIP OBJECTIVE

We can interpret the DIP objective in Eq. (3) in terms of the involved sets. If we neglect the noise term in Eq. (1), then the solution set is an affine space. Let us denote this space by $\mathbb{V}$ and let

$$\mathcal{T} = \{T_\theta(\boldsymbol{z}) | \theta \in \mathbb{R}^p\} \qquad (5)$$

denote the manifold described by the CNN. We define the $\boldsymbol{A}$-weighted distance $d_{\boldsymbol{A}}$ of a point $\boldsymbol{x}$ to a set $\mathcal{S}$ as

$$d_{\boldsymbol{A}}(\boldsymbol{x}, \mathcal{S}) = \min_{\boldsymbol{s} \in \mathcal{S}} \|\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{s})\|^2.$$

Hence, DIP essentially returns the point $\boldsymbol{x}^* \in \mathcal{T}$ that minimizes the $\boldsymbol{A}$-weighted distance to $\mathbb{V}$, i.e.

$$\boldsymbol{x}_{\mathrm{DIP}}^* = \arg\min_{\boldsymbol{x} \in \mathcal{T}} d_{\boldsymbol{A}}(\boldsymbol{x}, \mathbb{V}). \qquad (6)$$

However, constraining the solution to $\mathcal{T}$ has disadvantages. To begin with, $\mathbb{V}$ and $\mathcal{T}$ are usually disjoint, meaning that the result will never exactly fulfill the constraints posed by the inverse problem. For instance, a superresolved image $\boldsymbol{x}^*$ obtained from a downscaled version $\boldsymbol{y}$ via Eq. (6), will not result in $\boldsymbol{y}$ when downscaled again. While this may be desirable in problems involving noise, in cases like noiseless superresolution or inpainting, it is not. But even in cases, where noise is involved, it is unlikely that $\mathcal{T}$ is the ideal model of a natural image manifold. Another problem with Eq. (6) can arise when used in combination with regularizers to promote certain features of $\boldsymbol{x}$ that are not captured by $\mathcal{T}$. To illustrate this problem, consider the extreme case, where the employed regularizer $R(\boldsymbol{x})$ is such that $\mathcal{T}$ is a level set, i.e. $R(\boldsymbol{x}) = \text{const.} \ \forall \boldsymbol{x} \in \mathcal{T}$. Then, adding $R(\boldsymbol{x})$ to Eq. (6) has no effect on the optimization. A naive remedy to these problems is to search for an $\boldsymbol{x}_{\mathrm{btwn}}$ somewhere between $\mathcal{T}$ and $\mathbb{V}$
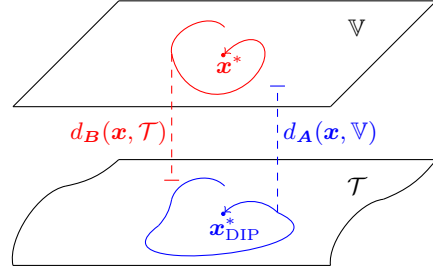

Fig. 1: Optimization procedures of DIP and SUB-DIP.

by means of the objective

$$\min_{\boldsymbol{x}_{\mathrm{btwn}} \in \mathbb{R}^n, \theta \in \mathbb{R}^p} \|T_\theta(\boldsymbol{z}) - \boldsymbol{x}_{\mathrm{btwn}}\|^2 + \lambda \|\boldsymbol{A}\boldsymbol{x}_{\mathrm{btwn}} - \boldsymbol{y}\|^2, \quad (7)$$

where $\lambda > 0$ is a tuning parameter. Unfortunately, we observed that this approach rarely has success due to local minima. Specifically, $\boldsymbol{x}_{\mathrm{btwn}}$ tends to quickly converge to early estimations of $T_\theta(\boldsymbol{z})$ and not evolve any further.

As an alternative, we propose an approach based on an inverted perspective of Eq. (6). Specifically, we aim to search for a point in $\mathbb{V}$ that minimizes some distance to $\mathcal{T}$, i.e., for an appropriate choice of some full rank matrix $\boldsymbol{B} \in \mathbb{R}^{m \times n}, m \leq n$, we formulate the objective

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x} \in \mathbb{V}} d_{\boldsymbol{B}}(\boldsymbol{x}, \mathcal{T}). \qquad (8)$$

The matrix $\boldsymbol{B}$ can be seen as an additional degree of freedom in the objective that allows us to measure the distance in a way such that the algorithm does not get stuck at early estimations of $\boldsymbol{x}$. Fig. 1 visualizes the optimization procedures of DIP and our subspace induced modification, that we refer to as *SUB-DIP* in the following. DIP is performing its optimization (blue curve) by searching for the point on $\mathcal{T}$ that minimizes the distance (red dashed line) to $\mathbb{V}$. In our approach, the optimization (red curve) is performed by searching for the point on $\mathbb{V}$ that minimizes the distance to $\mathcal{T}$. In principle, Eq. (8) could be solved by combining the classical DIP algorithm with appropriate projection operations. However, such an approach would not allow for an additional regularization of $\boldsymbol{x}$, which is why we prefer to optimize Eq. (8) directly. Eq. (8) does not exhibit the regularizer problem of DIP, since a regularizer with $\mathbb{V}$ as a level set contradicts the very definition of Eq. (1).

By construction, the solution in Eq. (8) exactly fulfills the constraint in Eq. (1) for $\boldsymbol{\eta} = 0$. This noiseless case is the main focus of our work. However, when only noisy observations are available, the solution $\boldsymbol{x}$ must depart from the subspace $\mathbb{V}$ To do so, we can modify Eq. (8) by incorporating a relaxation term $\boldsymbol{\xi} \in \mathbb{V}_\perp$, where $\mathbb{V}_\perp$ is the row space of $\boldsymbol{A}$, i.e. the orthogonal complement to the vector space that is parallel to $\mathbb{V}$. A relaxed version of Eq. (8) can be thus formulated as

$$\boldsymbol{x}^*, \boldsymbol{\xi}^* = \arg\min_{\boldsymbol{x} \in \mathbb{V}, \boldsymbol{\xi} \in \mathbb{V}_\perp} d_{\boldsymbol{B}}(\boldsymbol{x} + \boldsymbol{\xi}, \mathcal{T}) + \lambda \|\boldsymbol{\xi}\|^2, \qquad (9)$$

where $\lambda > 0$ is a tuning parameter that should be chosen inversely proportional to the noise variance.

The formulations in Eq. (8) and Eq. (9) allow us to better exploit the constraints in Eq. (1), either by limiting the solution set entirely to $\mathbb{V}$ or by penalizing the distance to $\mathbb{V}$.

Additionally to not having the disadvantages of the original DIP formulation as discussed in the beginning of this section, we observed that the formulations in Eq. (8) and Eq. (9) are less prone to numerical failures, when used in combination with intricate regularizers. This is due to the fact that unlike Eq. (6), x is not a function of θ and thus the gradient does not need to be backpropagated through several layers of the CNN. Finally, by choosing the weighting matrix B appropriately, we can avoid getting stuck in local minima as it is the case for Eq. (7).

4. APPLICATION TO SUPERRESOLUTION

In the following, we describe an algorithm that applies SUB-DIP to the problem of superresolution. For the sake of clarity, we derive the procedure on 1D signals but emphasize that the generalization to 2D signals is straightforward. Indeed, experiments are carried out for 2D images.

4.1. Superresolution Algorithm

Superresolution refers to the problem of reconstructing a high-resolution signal $x \in \mathbb{R}^{\tau d}$ from a low-resolution version $y \in \mathbb{R}^d$ with $\tau \in \mathbb{N}$ being the magnification factor. It is generally assumed that y results from subsequent filtering of $x \in \mathbb{R}^{\tau d}$ with a filter $h \in \mathbb{R}^L$, and downsampling by the factor τ i.e.

$$y = \mathrm{DS}_\tau[h * x], \tag{10}$$

where $\mathrm{DS}_\tau[\cdot]$ is a subsampling operator and $*$ denotes convolution[1]. Eq. (10) can be written in the form of Eq. (1). Without loss of generality, we assume that $L = k\tau, k \in \mathbb{N}$. The vector h can thus be split into k subvectors $h_i \in \mathbb{R}^\tau$ as follows.

$$h^\top = \begin{bmatrix} h_1^\top & h_2^\top & \cdots & h_{k-1}^\top & h_k^\top \end{bmatrix}. \tag{11}$$

Let us definite the matrix $H \in \mathbb{R}^{d \times \tau d}$ as

$$H = \begin{bmatrix} h_1^\top & \cdots & & h_k^\top & & & & \\ & h_1^\top & & \cdots & & h_k^\top & & \\ & & \ddots & & & & \ddots & \\ & & & h_1^\top & \cdots & & & h_k^\top \\ & & & & h_1^\top & \cdots & & h_{k-1}^\top \\ & & & & & \ddots & & \vdots \\ & & & & & & h_1^\top & h_2^\top \\ & & & & & & & h_1^\top \end{bmatrix}. \tag{12}$$

Then, Eq. (10) is equivalent to $y = Hx$. With $x_0 = H^\dagger y$, the solution set can be written as

$$\mathbb{V} = \{x_0 + G^\top s | s \in \mathbb{R}^{(\tau-1)d}\}, \tag{13}$$

[1]Mathematically, the operation is actually a cross-correlation

where G is a matrix with columns that span the kernel $\ker(H)$ of H. We can now rewrite Eq. (8) to fit our superresolution problem, as follows.

$$\min_{x \in \mathbb{V}} d_B(x, \mathcal{T}) = \min_{\substack{s \in \mathbb{R}^{(\tau-1)d}, \\ \theta \in \mathbb{R}^p}} \|B(x_0 + G^\top s - T_\theta(z))\|^2. \tag{14}$$

Similarly to Eq. (9), we can include noise into Eq. (14) via[2]

$$\min_{\substack{s \in \mathbb{R}^{(\tau-1)d} \\ \xi \in \mathbb{R}^d, \theta \in \mathbb{R}^p}} \|B(x_0 + G^\top s + H^\top \xi - T_\theta(z))\|^2 + \lambda_\xi \|\xi\|^2. \tag{15}$$

Equations (14) and (15) can be optimized by applying any common variation of gradient descent to θ and s. The optimization procedure of SUB-DIP is almost identical to the one of DIP, but includes a gradient step for s in each iteration.

The matrices B and G are extremely large for real-world signals and the backpropagation of the gradient is computationally only feasible, if they exhibit a convolutional or transpose-convolutional structure. In other words, we need to find a set of filters that can replace the respective matrix multiplications by convolutions.

For G, we need to make sure, that the chosen filters $g^1, \ldots, g^{\tau-1}$ span $\ker(H)$. Candidates for $g^1, \ldots, g^{\tau-1}$ can be generated in the following way. We consider the submatrix $H_{:,\tau(d-k)+1:\tau d}$ of H that contains only its last L columns. Note that only the last $2k-1$ rows of the matrix are non-zero. We can then generate an orthogonal basis $\gamma_1, \ldots, \gamma_{L+1-2k}$ of $\ker(H_{:,\tau(d-k)+1:\tau d})$. Then, for any matrix $\Gamma_i \in \mathbb{R}^{d \times \tau d}$ that is constructed from a basis element γ_i in the same manner as H is constructed from h, the rows lie in $\ker(H)$.

The matrix B is chosen to fulfill two properties. It should be easily realizable by convolutional operations and allows us to decompose Eq. (14) and Eq. (15) into its components on $\mathbb{V}$ and $\mathbb{V}_\perp$. We thus fix

$$B = \begin{bmatrix} H^\top & \sqrt{\lambda_G} G^\top \end{bmatrix}^\top, \tag{16}$$

where $\lambda_G > 0$ is a parameter. Since $HG^\top = 0$, we get

$$\begin{aligned} \|B(x_0 + G^\top s - T_\theta(z))\|^2 &= \|H(x_0 - T_\theta(z))\|^2 \\ &+ \lambda_G \|G(G^\top s - T_\theta(z))\|^2, \end{aligned} \tag{17}$$

for the Eq. (14) and a similar decomposition for Eq. (15). The two terms correspond to the known, and interpolated parts of x, respectively. Choosing an appropriate weighting factor λ_G permits us to avoid running into early local minima, as mentioned before. By putting less emphasis on the second term, we can ensure that $T_\theta(z)$ does not get optimized to approach meaningless early estimations of the unknown parts of x.

4.2. TV Regularization

Our new formulation, allows to easily introduce new regularizers. We illustrate this by considering TV. It was shown [7]

[2]$H^\top \xi$ is in $\mathbb{V}_\perp$, because $\mathbb{V}_\perp$ is the row space of H.

that TV regularization can further improve the performance of DIP. For a 2D image $\boldsymbol{I}$, a TV regularizer is formulated as

$$r_{\text{TV}}(\boldsymbol{I}) = \sum_{i,j} \sqrt{(\boldsymbol{I}_{i,j} - \boldsymbol{I}_{i-1,j})^2 + (\boldsymbol{I}_{i,j} - \boldsymbol{I}_{i,j-1})^2}. \quad (18)$$

With TV regularization Eq. (8) becomes

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x} \in \mathbb{V}} d_{\boldsymbol{B}}(\boldsymbol{x}, \mathcal{T}) + \lambda_{\text{TV}} r_{\text{TV}}(\boldsymbol{x}). \quad (19)$$

5. EXPERIMENTS

We evaluate our algorithm on the 4x superresolution of the Set5 [11] dataset. To this end, we calculate the PSNR values for the superresolved images reconstructed by our algorithm. The emphasis of this section is to perform an ablation study that compares the original DIP to SUB-DIP. We reuse the official DIP implementation that was made publicly available [12]. We directly optimize the objectives in Eq. (14) and Eq. (15) by means of the Adam optimizer and fix $\lambda_{\boldsymbol{G}} = 0.1$. The same parameters for DIP as for SUB-DIP are used. In particular, a Lanczos filter is used as $\boldsymbol{h}$, even though the actual filter that was used to perform the downsampling is unknown to us. Since many superresolution algorithms operate on grayscale images only, it is common to calculate the PSNR exclusively on the luminance channel of an image. Here, we are more interested in how SUB-DIP compares to other RGB algorithms, in particular the original DIP formulation, which is why we compute the PSNR on all three RGB channels. This explains why the values reported in [5], are slightly higher. DIP performs some cropping in order to create images with dimensions divisible by 32. For the sake of comparability, we thus crop the ground truth images accordingly, as well as the images generated by all other baseline algorithms, obtained from [12].

5.1. Noiseless Superresolution

Often, superresolution does not involve any noise. This is for instance the case when the downsampling was done digitally. In that case, we can search for the solution directly on $\mathbb{V}$ by means of Eq. (14). We evaluate DIP and SUB-DIP both without any regularization as well as with an additional TV term as described in Section 4.2, with $\lambda_{TV} = 10^{-3}/n$. Table 1 shows the reconstruction results.

SUB-DIP consistently outperforms DIP without TV regularization. With an added TV term, DIP is also outperformed by SUB-DIP, except for the "Woman" image. The advantage of our formulation of the DIP objective can be observed in Fig. 2. The manifold $\mathcal{T}$ tends not to capture certain directional stuctures at very high frequencies. Restricting the solution to $\mathcal{T}$ thus causes a blurring out of such structures. This is why, for instance, the DIP struggles to reproduce fine details such as eyelashes, while SUB-DIP does not have this

	Baby	Bird	Butterfly	Head	Woman	Avg.
Bicubic	30.43	28.09	20.90	28.72	25.39	26.71
LapSRN	32.02	30.42	25.52	29.62	29.24	29.36
DIP	29.85	29.62	24.83	28.49	27.00	27.96
DIP+TV	29.87	29.52	24.60	28.62	27.20	27.96
Ours	31.38	29.99	24.84	28.74	27.01	28.39
Ours+TV	31.47	29.98	24.81	28.67	26.85	28.35

Table 1: PSNR values for Set5. Best and second-best results are written in red, and blue respectively.
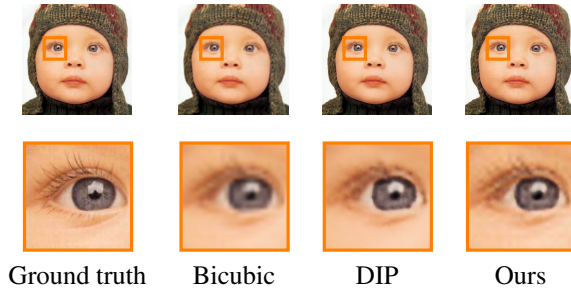

| Ground truth | Bicubic | DIP | Ours |

Fig. 2: Reconstruction results for "Baby"

problem to that extent. State-of-the-art deep learning methods [13, 14, 15] such as LapSRN [3] achieve generally better performance than DIP or SUB-DIP, but require training on large datasets, which is not always feasible.

5.2. Noisy Superresolution

We test the approach in Eq. (15) by adding Gaussion noise with $\sigma_{noise} = 0.02 * I_{\max}$ to the downsampled image $\boldsymbol{y}$, where $I_{\max}$ corresponds to the maximal pixel value. The noise regulization weight is set to $\lambda_{\boldsymbol{\xi}} = 3 * 10^{-4}/(d * \sigma_{noise}^2)$. Even though our method can no longer exploit the subspace assumption, it still achieves a slight improvement over classical DIP, as can be seen in Table 2.

	Baby	Bird	Butterfly	Head	Woman	Avg.
DIP	29.24	27.70	23.81	26.95	26.17	26.78
Ours	29.74	28.22	23.89	27.47	26.40	27.15

Table 2: PSNR values for Set 5 with noise

6. CONCLUSION

In this work, we presented a novel approach to leverage the Deep Image Prior by formulating an optimization procedure on the solution set of an inverse problem. We have described how to apply this approach on the problem of superresolution and have demonstrated a significant improvement of the reconstruction results.

7. REFERENCES

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

[2] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma, "Image super-resolution as sparse representation of raw image patches," in *2008 IEEE conference on computer vision and pattern recognition*. IEEE, 2008, pp. 1–8.

[3] Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang, "Deep laplacian pyramid networks for fast and accurate super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 624–632.

[4] Raymond A Yeh, Chen Chen, Teck Yian Lim, Alexander G Schwing, Mark Hasegawa-Johnson, and Minh N Do, "Semantic image inpainting with deep generative models," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5485–5493.

[5] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky, "Deep image prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9446–9454.

[6] Sören Dittmer, Tobias Kluth, Peter Maass, and Daniel Otero Baguer, "Regularization by architecture: A deep prior approach for inverse problems," *arXiv preprint arXiv:1812.03889*, 2018.

[7] Jiaming Liu, Yu Sun, Xiaojian Xu, and Ulugbek S Kamilov, "Image restoration using total variation regularized deep image prior," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 7715–7719.

[8] Dave Van Veen, Ajil Jalal, Mahdi Soltanolkotabi, Eric Price, Sriram Vishwanath, and Alexandros G Dimakis, "Compressed sensing with deep image prior and learned regularization," *arXiv preprint arXiv:1806.06438*, 2018.

[9] Gary Mataev, Michael Elad, and Peyman Milanfar, "Deepred: Deep image prior powered by red," *arXiv preprint arXiv:1903.10176*, 2019.

[10] Yaniv Romano, Michael Elad, and Peyman Milanfar, "The little engine that could: Regularization by denoising (red)," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1804–1844, 2017.

[11] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proceeding of the British machine vision Conference*. 2012, BMVA press.

[12] Dmitry Ulyanov, "Deep image prior: Project website," `https://dmitryulyanov.github.io/deep_image_prior`, 2008.

[13] Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita, "Deep back-projection networks for super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1664–1673.

[14] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1646–1654.

[15] Mehdi SM Sajjadi, Bernhard Scholkopf, and Michael Hirsch, "Enhancenet: Single image super-resolution through automated texture synthesis," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 4491–4500.