# Single-Shot Plug-and-Play Methods for Inverse Problems

**Anonymous authors**
**Paper under double-blind review**

## Abstract

The utilisation of Plug-and-Play (PnP) priors in inverse problems has become increasingly prominent in recent years. This preference is based on the mathematical equivalence between the general proximal operator and the regularised denoiser, facilitating the adaptation of various off-the-shelf denoiser priors to a wide range of inverse problems. However, existing PnP models predominantly rely on pre-trained denoisers using large datasets. In this work, we introduce Single-Shot PnP methods (SS-PnP), shifting the focus to solving inverse problems with minimal data. First, we integrate Single-Shot proximal denoisers into iterative methods, enabling training with single instances. Second, we propose implicit neural priors based on a novel function that preserves relevant frequencies to capture fine details while avoiding the issue of vanishing gradients. We demonstrate, through extensive numerical and visual experiments, that our method leads to better approximations.

## 1 Introduction

Inverse problems have long been a fundamental challenge in the field of mathematics and applied sciences, encompassing a wide range of applications from image reconstruction to signal processing Devaney (2012); Bertero et al. (2021). Traditionally, these problems have been approached through various analytical and numerical methods Vogel (2002); Jordan (1881); Metropolis & Ulam (1949). However, the advent of deep learning has revolutionised this domain. Deep inverse problems present a modern approach, offering new insights and solutions where conventional methods have limitations McCann et al. (2017). This shift towards leveraging machine learning techniques marks a significant evolution in tackling inverse problems, opening doors to more sophisticated and efficient problem-solving techniques.

A popular framework in this deep inverse problem era is Plug-and-Play (PnP) methods Venkatakrishnan et al. (2013); Zhang et al. (2021); Chan et al. (2016); Ono (2017). At the core of this approach lies the mathematical equivalence of the proximal operator to denoising Venkatakrishnan et al. (2013), a concept that intertwines optimisation theory with modern denoisers. This equivalence paves the way for the integration of advanced deep learning-based denoisers into the inverse problem-solving process. The Plug-and-Play framework essentially allows for the seamless insertion of these denoisers into iterative algorithms, enhancing their ability to recover high-quality signals or images from corrupted observations Zhang et al. (2021; 2019); Ahmad et al. (2020).

Although PnP methods have demonstrated outstanding results across a wide range of inverse problems, they predominantly rely on the assumption of having substantial training datasets for the development of robust denoising models Arridge et al. (2019). This requirement often becomes a significant bottleneck, especially in scenarios where data availability is limited or the diversity of data is not sufficiently representative. Moreover, PnP methods typically necessitate retraining or fine-tuning to adapt effectively to varying distributions and signal characteristics. To address these challenges, Single-Shot learning emerges as a promising alternative, offering a paradigm shift in how deep inverse problems can be trained with minimal data. Unlike traditional methods that require extensive datasets, Single-Shot learning aims to make significant inferences from a single instance, or in some cases, a small set of instances. *To the best of our knowledge, there is no existing work on Single-Shot PnP methods. Our work thus opens the door to a novel research line for PnP, introducing the concept of Single-Shot Plug-and-Play methods.*

Single-Shot techniques offer a viable solution to the current constraints in PnP methods, particularly in their reliance on extensive training datasets. This shift enables us to delve into the recent advancements in signal representation, specifically, the emergence of neural implicit representations. Neural implicit representations Strümpler et al. (2022); Saragadam et al. (2023); Tancik et al. (2020) are trained to learn continuous functions that map coordinates to signal values or features, making this form of representation highly efficient in capturing intricate data details with a compact network architecture. Its application in the context of Single-Shot PnP methods is particularly promising, as denoisers that can effectively operate with limited data inputs, aligning with the Single-Shot learning paradigm. *We therefore use implicit neural representations as a way to represent the proximal denoiser in PnP techniques in a Single-Shot fashion.* Our contributions are:

- We propose the concept of Single-Shot Plug-and-Play (SS-PnP) Methods to solve inverse problems, in which we highlight:
  - We introduce Single-Shot proximal denoisers into iterative methods for solving any inverse problem. Our scheme eliminates the need for a pre-trained model, enabling training with a single instance.
  - We introduce implicit neural priors for Plug-and-Play methods that enables the network to preserve more details during training. Additionally, we provide a theoretical justification for our prior, emphasising how its continuity and differentiability play a crucial role in mitigating the issue of vanishing gradients during the training process and preserve fine details.
- We demonstrate, through extensive experiments on several inverse problems, that our technique leads to a better approximation on capturing finer details, smoother edge features and better colour representation. The method is evaluated on inverse problems with both single operator task and multi-operator task like joint demosaicing and deconvolution task. With only one image input in the whole reconstruction process, it outperforms the classical methods and pre-trained models among all the tasks.

## 2 Related Work

In this section, we review the existing literature and the concepts closely related to our work.

**Plug-and-Play (PnP) Methods.** They have revolutionised the field of inverse problems by integrating advanced denoisers into iterative algorithms. This innovative approach, initiated by Venkatakrishnan et al. Venkatakrishnan et al. (2013), has undergone significant evolution. Meinhardt et al. Meinhardt et al. (2017) showcased its effectiveness in diverse imaging tasks, marking a turning point in PnP's development. The subsequent works Ryu et al. (2019); Sun et al. (2019); Hurault et al. (2022) further refined PnP, enhancing its stability and convergence, thus broadening its applicability.

In addition, the works of that Teodoro et al. (2018); Yuan et al. (2020); Zhang et al. (2017b); Ono (2017); Sun et al. (2019) further expanded the scope of PnP, demonstrating its adaptability in complex imaging scenarios. A notable advancement in the optimisation landscape came with the introduction of TFPnP (Tuning-Free Plug-and-Play) Wei et al. (2020), which innovatively eliminated the need for parameter tuning in PnP algorithms.

Previous methods have relied on data-driven pre-training, which becomes impractical in situations with limited data or on smaller devices due to the extensive size of the required models. Consequently, developing a method for Single-Shot image prior denoising emerges as a compelling solution for such resource-constrained environments.

**Single-Shot Image Denoising.** A crucial element in PnP methods is the denoiser model. During the last years, the denoiser technique in PnP has evolved remarkably, transitioning from traditional methods to advanced deep learning techniques. Pioneering works such as the BM3D algorithm Dabov et al. (2007) and the Non-Local Means (NLM) algorithm Buades et al. (2005) laid the groundwork, setting significant benchmarks.

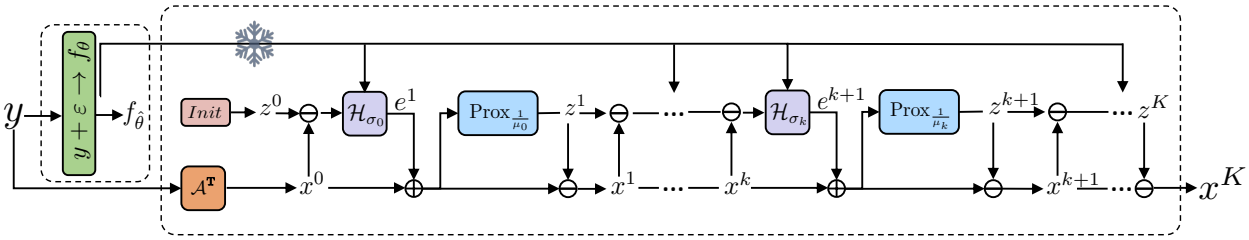


Figure 1: The pipeline of Single-Shot Plug-and-Play methods (SS-PnP). The 2 blocks indicate the 2 steps as in Algorithm 1, and the ❄ indicates the fix in denoiser weight over the ADMM iterations, where there are $K$ iterations in total (i.e. $k \in \{0, 1, ..., K-1\}$).

The introduction of deep learning marked a paradigm shift, exemplified by the DnCNN model proposed by Zhang et al. Zhang et al. (2017a) and its variant DnCNN-S Zhang et al. (2018), which demonstrated the efficacy of convolutional networks in denoising. The Deep Image Prior (DIP) by Ulyanov et al. Lempitsky et al. (2018) furthered this progression, utilising the structure of convolutional networks as a prior for denoising.

A notable advancement is the rising of self-supervised methods, which revolutionised the Single-Shot denoising field by eliminating the need for clean training data. The Noise2Void framework by Krull et al. Krull et al. (2019) and the Noise2Self algorithm Batson & Royer (2019) are pioneering examples, utilising concepts like blind-spot networks for effective denoising. Building on these, Self2Self Quan et al. (2020), Noise2Same Xie et al. (2020), and Noise2Info Wang et al. (2023) further explored self-supervision, offering unique strategies for leveraging the inherent properties of noisy images. Additional approaches like CycleISP Zamir et al. (2020), IRCNN Zhang et al. (2017b), GCDN Anwar & Barnes (2020), and bayesian denoising with blind-spot networks Laine et al. (2019) further enrich the landscape, each contributing novel perspectives and solutions to the challenge of Single-Shot image denoising.

These advancements are not confined to noise reduction alone; many of the developed deep denoisers are inherently adaptable and can be generalised to tackle various inverse problems in imaging. Inverse problems, such as demosaicing, denoising, and deconvolution, share common traits with denoising. The underlying principles and network architectures developed for denoising can often be extended or fine-tuned to address these challenges Romano et al. (2017); Akyüz et al. (2020). Furthermore, the concept of Plug-and-Play methods opens up new avenues. *However, this progress highlights a gap: despite the evolution of Single-Shot denoisers, there is currently no work on Single-Shot denoisers into iterative algorithms. Therefore, this work introduces the concept of Single-Shot Plug-and-Play methods.*

Despite the versatility of Single-Shot image denoising, CNN-based estimators frequently fail to capture continuously high-frequency details crucial for image reconstruction. Implicit neural representation (INR) emerges as a solution, adept at addressing these high-frequency challenges in inverse problems.

**Implicit Neural Representation**, characterised as a fully connected network-based method, has seen a rise in popularity for solving inverse problems as highlighted by Sun et al. Sun et al. (2021). Traditional activation functions like ReLU have exhibited limitations in representing high-frequency features, as discussed by Dabov et al. Dabov et al. (2007). This shortcoming has led to the exploration of nonlinear activation functions, such as the sinusoidal function Sitzmann et al. (2020), enhancing representational capabilities. The adaptability of INR is evident in its diverse applications across medical imaging Wang et al. (2022), image processing Chen et al. (2021); Attal et al. (2021), and super-resolution Saragadam et al. (2023).

A distinct advantage of INR lies in its independence from pre-training, attributed to its rapid training capability, as indicated by Saragadam et al. Saragadam et al. (2023). This feature makes INR particularly suitable for Single-Shot image denoising tasks, where a single image suffices for effective learning, bypassing the need for extensive data-driven pre-training. The ability of INR to learn from minimal data points underscores its potential in resource-limited scenarios.

## 3 Methodology

Inverse imaging problems have emerged as a crucial cornerstone in computer vision and various related domains. A conventional forward model reads:

$$y = A(x) + \epsilon, \tag{1}$$

where $y \in \mathbb{R}^M$ is the known observation, $x \in \mathbb{R}^N$ is the unknown target of interest, $\epsilon \in \mathbb{R}^M$ is the measurement noise, and $A : \mathbb{R}^N \to \mathbb{R}^M$ is the forward measurement operator, which varies across different tasks.

The approximation of $x$ gives rise to a highly ill-posed problem, necessitating regularisation. Consequently, the fundamental objective for any imaging inverse problem is the minimisation of:

$$\min_{x \in \mathbb{R}^N} D(x) + \gamma R(x), \tag{2}$$

$D(x)$ refers to the data fidelity term, usually taking the form $D(x) = \frac{1}{2}\|A(x) - y\|^2$. $R(x)$ denotes the regularisation term, which encodes prior knowledge about $x$. The parameter $\gamma$ serves as the regularisation weight, determining the trade-off between data fidelity and regularisation.

### 3.1 Single-Shot Proximal Denoiser

We introduce Single-Shot Plug-and-Play methods (SS-PnP) as demonstrated in Figure 1. The optimisation of equation 2 typically exhibits non-smooth characteristics due to $R$. A widely adopted strategy for addressing this problem is to employ first-order methods such as the alternating direction method of multipliers (ADMM). Given a function $F(\cdot)$, we define the proximal operator of $F$ at $v$ with step size $\delta$ as:

$$\mathrm{Prox}_{\delta F}(v) = \underset{u}{\mathrm{argmin}} \left\{ \frac{1}{2}\|u - v\|^2 + \delta F(u) \right\}, \tag{3}$$

Considering ADMM, one can express PnP-ADMM as:

$$\begin{aligned} \bigstar \quad e^{k+1} &= \mathrm{Prox}_{\sigma_k^2 R}\left(z^k - x^k\right) = \mathcal{H}_{\sigma_k}(z^k - x^k) \\ z^{k+1} &= \mathrm{Prox}_{\frac{1}{\mu_k} D}\left(e^{k+1} + x^k\right) \\ x^{k+1} &= x^k + e^{k+1} - z^{k+1} \end{aligned} \tag{4}$$

where $\mathrm{Prox}_{\sigma_k^2 R}(\cdot)$ is the proximal operator of the regularisation with noise strength $\sigma_k$ and $\mathrm{Prox}_{\frac{1}{\mu_k} D}(\cdot)$ is to enforce the data consistency Ryu et al. (2019) with penalty parameter $\mu$ in the $k$-th iteration, for $k \in \{0, 1, 2, ..., K-1\}$. From equation 4 -$\bigstar$, we can observe that Plug-and-Play (PnP) methods leverage the equivalence between the proximal operator $\mathrm{Prox}_{\sigma_k^2 R}(\cdot)$ and a denoiser $\mathcal{H}_{\sigma_k}$ with the denoising parameter $\sigma_k \geq 0$.

**Single-Shot Denoiser is All You Need for PnP.** PnP techniques primarily rely on denoisers, often trained on extensive datasets and using off-the-shelf deep denoisers Ryu et al. (2019); Sun et al. (2019); Hurault et al. (2022). However, an unexplored question is *whether Single-Shot denoisers can be used in iterative methods and what their properties are. To our knowledge, there are no existing works on this.*

In our Single-Shot denoising stage, we utilise the observed image with noise $y + \varepsilon$ to train the denoiser, enabling it to distinguish and mitigate complex noise and distortions specific to the example. The neural network $f_\theta$ aims to transform the noisy and corrupted image into its less corrupted counterpart. This is achieved through an optimisation process given by:

$$\hat{\theta} = \underset{\theta}{\mathrm{argmin}}\, \mathcal{L}(f_\theta, y + \varepsilon), \tag{5}$$

where $\mathcal{L}$ is a loss function that evaluates the difference between the network's output and the corresponding observed image with noise. Refer to Step 1 in Algorithm 1.

**Algorithm 1:** Single-Shot Plug-and-Play

**Forward model:** $y = A(x) + \epsilon$
**Input:** $y \in \mathbb{R}^M$
**Output:** $x \in \mathbb{R}^N$

**Step 1** Train the Single-Shot Denoiser

Choose noise strength $\varepsilon$, and initialise $f_{\theta_0}$ ▸ $\varepsilon$ is independent of $\epsilon$
**for** $i \in \{0, 1, ..., I-1\}$ **do**
  $f_{\theta_{i+1}} \leftarrow \mathcal{L}(f_{\theta_i}, y + \varepsilon)$ ▸ Train the Single-Shot denoising prior
**end**
$f_{\hat{\theta}} = f_{\theta_I}$

**Step 2** ADMM Iteration

Let $\mathcal{H} = f_{\hat{\theta}}$ , and $x^0 = A^T(y)$ ▸ $A^T$ denotes adjoint of $A$
Choose noise strength $\sigma : \{\sigma_k\}_0^{K-1}$ , and penalty parameter $\mu : \{\mu_k\}_0^{K-1}$
Initialise $z^0$
**for** $k \in \{0, 1, 2, ..., K-1\}$ **do**
  $e^{k+1} = \mathcal{H}_{\sigma_k}(z^k - x^k)$
  $z^{k+1} = \mathrm{Prox}_{\frac{1}{\mu_k} D}\left(e^{k+1} + x^k\right)$
  $x^{k+1} = x^k + e^{k+1} - z^{k+1}$
**end**
$x = x^K$

This pre-trained model is subsequently integrated into a Plug-and-Play (PnP) framework as a prior for regularisation. Let $\mathcal{H} = f_{\hat{\theta}}$, the trained denoiser serves as a guiding force in the iterative reconstruction process, enhancing the ability to recover high-quality images from corrupted observations. By embedding this Single-Shot learning model into the PnP framework, we establish a potent approach for tackling a range of inverse problems, particularly in situations involving severely corrupted images. To promote the synergy between Single-Shot learning and PnP methods, the $f_\theta$ utilise nonlinear INR denoiser. Refer to Step 2 in Algorithm 1.

In this work, we consider ADMM justified by its well-known outstanding performance Venkatakrishnan et al. (2013); Wei et al. (2022) in comparison to existing methods, its well-established convergence properties Ryu et al. (2019), and its modular approach, which collectively ensure its effectiveness across various inverse problem scenarios.

### 3.2 Implicit Neural Prior for Plug-and-Play

Implicit neural representations enable the learning of continuous functions from a signal. State-of-the-art performance relies on deep denoisers like UNet Ronneberger et al. (2015) and FFDNet Zhang et al. (2018). While convolutional-based denoisers have shown impressive results, there is currently no research exploring the use of implicit neural representation (INR) as a prior for PnP methods. *INR can be used as a Single-Shot approach in equation 4, we then open the door to a new research direction for PnP.*

INR boasts remarkable expressive power and inductive bias, paving the way for innovative denoising approaches. Another significant contribution of this work is the introduction of a new Single-Shot framework for PnP. In particular, *we present a novel INR prior.* Unlike the majority of existing works, we provide a solid theoretical justification for its properties and behavior.

**Zooming into our Prior.** We next define our proposed implicit neural representation prior.

**Our INR Prior**

We define a nonlinear activation function and form our implicit neural representation, which reads:

$$\Phi(x) = \exp\{-(a_1 x + b_1)^2\}\sin(a_2 x + b_2) + \frac{1}{\exp\{-(a_1 x + b_1)\} + 1} \tag{6}$$

Our proposed activation function is nonlinear. Moreover, it has key properties of differentiability and continuity.

*Proof.* Let $g(x) = \exp\{-(a_1 x + b_1)^2\}$ and $f(x) = \sin(a_2 x + b_2)$. For $\forall c \in \mathbb{R}$,

$$\begin{aligned} &\frac{(fg)(x) - (fg)(c)}{x - c} \\ &= \frac{f(x)g(x) - f(c)g(x) + f(c)g(x) - f(c)g(c)}{x - c} \\ &= \frac{f(x) - f(c)}{x - c} g(x) + f(c)\frac{g(x) - g(c)}{x - c} \end{aligned} \tag{7}$$

Consequently,

$$\begin{aligned} &\lim_{x \to c} \frac{(fg)(x) - (fg)(c)}{x - c} \\ &= \lim_{x \to c} \left[ \frac{f(x) - f(c)}{x - c} g(x) + f(c)\frac{g(x) - g(c)}{x - c} \right] \\ &= \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \lim_{x \to c} g(x) + f(c) \lim_{x \to c} \frac{g(x) - g(c)}{x - c} \end{aligned} \tag{8}$$

Because both $g(x)$ and $f(x)$ are continuous and differentiable, $\lim_{x \to c} \frac{(fg)(x) - (fg)(c)}{x - c}$ exists for all $c \in \mathbb{R}$. Hence, $\exp\{-(a_1 x + b_1)^2\}\sin(a_2 x + b_2)$ is continuous and differentiable for all $c \in \mathbb{R}$. Consequently, the entire activation function is the sum of two differentiable functions, making it both continuous and differentiable. □

*Why are these properties interesting?* The nonlinearity, continuity, and differentiability brought by our formula make it more representative compared to traditional activation functions, enabling the network to preserve more details during training. Additionally, the continuity and differentiability of our formula play a crucial role in mitigating the issue of vanishing gradients during training.

**Convergence of our Prior**

Consider the function $\Phi(x)$ in equation 6, then it holds that:

$$\Phi(x) = \begin{cases} 0 & x \to -\infty \\ 1 & x \to \infty \end{cases} \tag{9}$$

This behavior ensures that $\Phi(x)$ remains bounded and avoids gradient explosion, making it a stable function for optimisation.

*Proof.* For $y > 0, \forall \epsilon > 0, \quad \exists \bar{y} = \sqrt{\max\left(0, \ln \frac{1}{\epsilon}\right)}$ when $y > \bar{y}$ ,

$$\begin{aligned} &\exp\left(y^2\right)\sin(a_2 x + b_2) < \exp\left(-y^2\right) \\ &= \exp\left\{-\max\left(0, \ln \frac{1}{\epsilon}\right)\right\} \\ &= \exp\{\min(0, \ln \epsilon)\} \\ &\leqslant \exp(\ln \epsilon) = \epsilon \end{aligned} \tag{10}$$

Table 1: The performance (PSNR(dB)) comparison of the 4 Single-Shot deep denoising priors (DnCNN, FFDNet, UNet and our proposed ★) on super resolution (SR) task with 2× and 4× upscalings in the Plug-and-Play framework.

| SS-PnP | Fractal | | Wolf | | Dog | | Peacock | | Tiger | | Bird | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (Backbone) | 2× | 4× | 2× | 4× | 2× | 4× | 2× | 4× | 2× | 4× | 2× | 4× |
| DnCNN | 19.29 | 17.45 | 20.71 | 17.16 | 18.29 | 14.31 | 19.03 | 16.79 | 19.22 | 14.28 | 21.90 | 18.68 |
| FFDNet | 22.04 | 19.59 | 21.12 | 22.85 | 24.02 | 21.98 | 22.31 | 18.99 | 23.02 | 19.11 | 26.34 | 21.86 |
| UNet | 18.58 | 14.18 | 21.78 | 15.83 | 18.71 | 14.34 | 19.56 | 16.25 | 18.87 | 15.52 | 21.82 | 16.53 |
| ★ | **25.43** | **21.26** | **27.21** | **23.30** | **25.80** | **22.00** | **22.46** | **19.27** | **23.12** | **19.45** | **26.44** | **22.06** |

Meanwhile, the function $\frac{1}{\exp(-y)+1}$ converges to 1 when $x \to \infty$. Consequently, the function is convergent to 1 when $x \to \infty$. □

By leveraging the convergence and bounded nature of the activation function, we can significantly mitigate the issue of exploding gradients. Moreover, as the model converges to 0, it also effectively reduces the impact of outliers, enhancing the overall robustness of our model.

## 4 Experiment

In this section, we describe the experiments undertaken to validate our proposed Single-Shot Plug-and-Play framework.

### 4.1 Experimental Setting

In our image preprocessing, we utilised dual resizing strategies. We sourcing the images with Creative Commons Licenses and resized to $512 \times 384$, and in the meantime tested on the selected data in (Bevilacqua et al., 2012; Zeyde et al., 2012) without resizing. We remind to the reader, the experiments on Single-Shot Plug-and-Play methods (SS-PnP) considered only a single image input in the whole pipeline.

**Training Scheme.** During the initial implicit neural representation (INR) pre-training phase, Gaussian noise with a standard deviation of 0.1 was introduced. The training was conducted over 100 iterations, with a network configuration comprising 2 hidden layers and 64 features per layer. The learning rate was set to 0.001. We then reconstruct the image for 5 ADMM iteration steps with dynamic noise strength and penalty parameter chosen by logarithmic descent over steps.

**Evaluation Protocol.** For comparative analysis, the Noise2Self pre-training scheme Krull et al. (2019), was applied to three networks, each undergoes 100 training iterations with a learning rate of 0.01. The DnCNN Zhang et al. (2017a) and FFDNet Zhang et al. (2018) architectures, were configured with 8 hidden layers and 64 feature maps per layer. Conversely, the UNet Ronneberger et al. (2015) architecture, employed a 4-times downsampling strategy and initiated with 32 feature maps in first convolutional operation. For a fair comparison, we fixed the setting for the optimisation step as used in our proposed strategy.

The empirical studies are trained and tested on NVIDIA A10 GPU with 24GB RAM. The Plug-and-Play phase ensued with the $\triangledown$-Prox toolbox Lai et al. (2023), adopting the default settings for all tasks. For evaluating our methods, we employed Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Higher PSNR and SSIM scores signal superior reconstruction quality.

### 4.2 Results & Discussion

This section shows all the numerical and visual results that support our method.

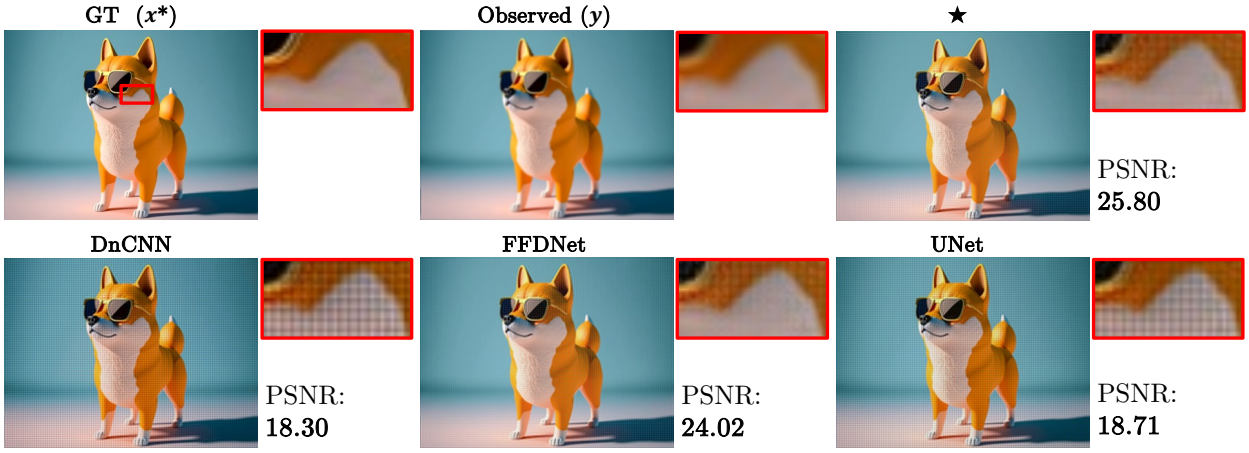


Figure 2: The comparative visualisation of super resolution with 2× upscaling task on "Dog" example among the 4 Single-Shot deep denoising priors (with ★ denotes our proposed prior) in Plug-and-Play framework. The zommed-in view is provided at the right hand side of each result.

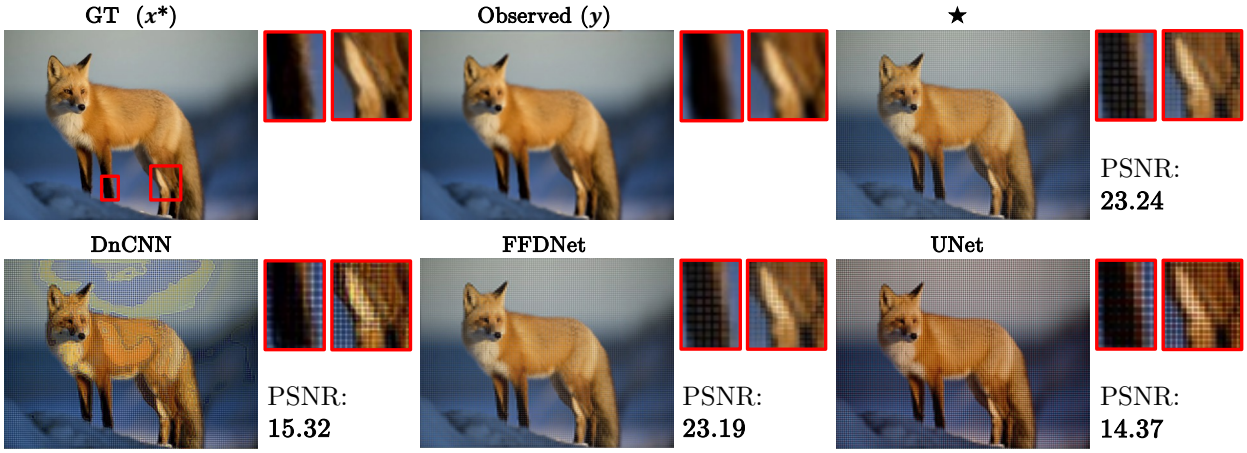


Figure 3: The visualisation comparison of super resolution with 4× upscaling task on "Fox" example between the 4 Single-Shot deep denoising priors performing within Plug-and-Play framework, with detailed comparison zoomed in. ★ denotes our prior.

**Super-resolution (SR).** It refers to the process of reconstructing a high-resolution image (HR) from one or more low-resolution observations (LR). The forward measurement operator for super-resolution is given by: $A(x) = (k \otimes x) \downarrow_s$, where $k$ represents the kernel for convoluting the image, and $\downarrow_s$ is the downscaling operation with scale $s$. Here, kernel operator was set as size 5 with standard deviation of the Gaussian distribution as 3.

In evaluating the Single-Shot super-resolution efficacy of our method, we utilise six varied categories of images, with upscaling factors of 2× and 4× to ensure a fair comparison against well-established techniques such as DnCNN, FFDNet, and UNet, which are all trained with the Noise2Self pre-training scheme. The quantitative results, as detailed in Table 1, reveal our method's outstanding performance, outshining the benchmarks with significant margins. Notably, our approach demonstrates a marked improvement in PSNR values, with particularly pronounced enhancements in the 4x upscaling scenario. These advancements indicate our model's robustness and substantial improvement forward in Single-Shot super resolution.

Our method demonstrates exceptional preservation of texture and detail complexity, as shown in Figures 2, 3. Our technique achieves high-resolution enhancement while intricately reconstructing fine details, evident from

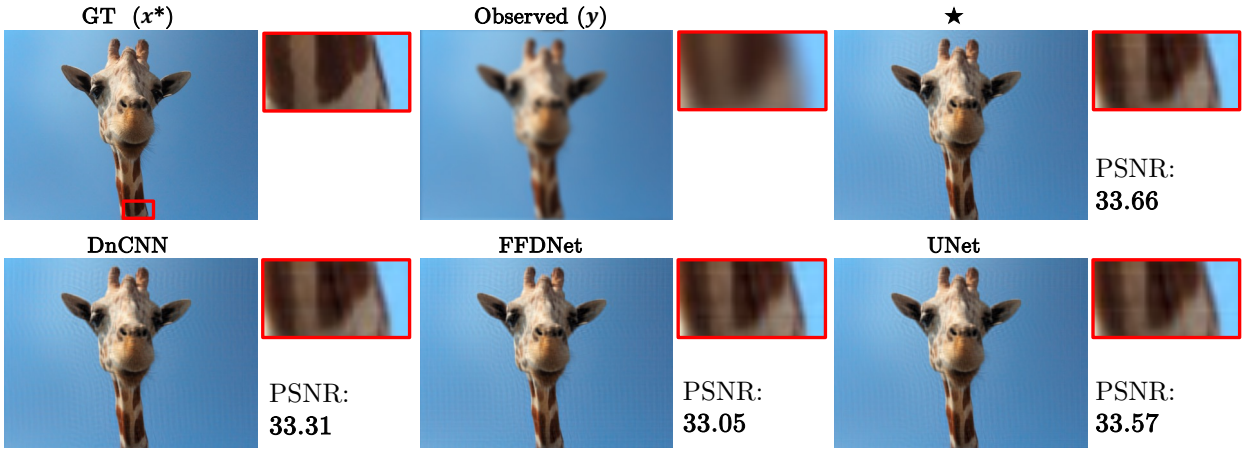


Figure 4: Visual comparison of the 4 deep denoising priors (include our proposed as ★) on deconvolution task in Single-Shot Plug-and-Play (SS-PnP) strategy for "Giraffe" example, with a zoomed-in region exhibiting intricate details.

Table 2: Evaluation comparison of the 4 Single-Shot deep denoising priors within the Plug-and-Play framework (SS-PnP setting) on deconvolution task measured with PSNR(dB), and SSIM. We denote our proposed Single-Shot denoising prior as ★.

| SS-PnP | Racoon | | Turtle | | Tiger | | Bird | | Head | | Monarch | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (Backbone) | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| DnCNN | 27.43 | 0.88 | 26.24 | 0.89 | 23.75 | **0.84** | 30.67 | 0.93 | 29.87 | **0.87** | 28.26 | 0.93 |
| FFDNet | 19.81 | 0.49 | 25.03 | 0.84 | 19.62 | 0.62 | 16.72 | 0.55 | 26.50 | 0.72 | 28.09 | 0.92 |
| UNet | 27.06 | 0.88 | 26.17 | **0.90** | 23.72 | **0.84** | 31.06 | **0.94** | 29.90 | **0.87** | 28.49 | **0.94** |
| ★ | **27.56** | **0.89** | **26.30** | **0.90** | **23.76** | **0.84** | **31.07** | **0.94** | **29.92** | **0.87** | **28.55** | **0.94** |

the minimised artifacts such as grid patterns and spurious spots in Figure 2. In Figure 3, DnCNN introduces noticeable colour distortions in both the background and the foreground. The UNet generates numerous undesirable grids. Although our approach presents a comparable visual quality to FFDNet, it achieves a superior PSNR, suggesting a quantitatively and qualitatively improved performance. Collectively, our results indicate that our method not only holds promise for practical application but also sets a new standard for super-resolution tasks.

**Image Deconvolution.** This is a computational technique aimed at reversing the effects of blur on photographs. Mathematically, the observed image, $y$, is the result of convoluing the true image, $x$, with a forward measurement operator: $A(x) = k \otimes x$. The goal of deconvolution is to estimate the original image $x$ by deconvoluting the observed image $y$ with the gaussian kernel $k$. Here, the kernel size was set as 15 with standard deviation 5 of Gaussian distribution.

The performance comparison of various Single-Shot deep denoiser algorithms on a deconvolution task, shown in Table 2, evaluated using both PSNR and SSIM metrics as well. The comparison includes our method alongside established techniques like DnCNN, FFDNet, and UNet across other six image categories. Our approach consistently achieves competitive PSNR scores, surpassing others in the 'Turtle' and 'Monarch' categories, and still shows parity improvements in SSIM values. This indicates not only enhanced accuracy in image reconstruction but also improved perceptual quality. These results underscore our method's effectiveness in noise reduction and sharpness, demonstrating its potential for practical deconvolution applications.

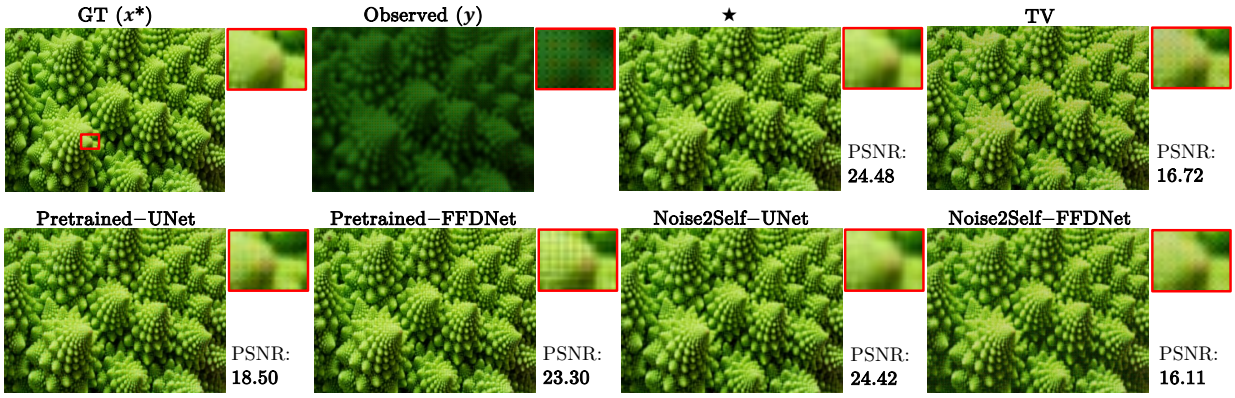


Figure 5: The comparative visualisation of joint deconvolution and demosaicing tasks on the "Fractal" example. The comparison is made across different denoising priors in Play-and-Play framworks, inclusing Single-Shot deep denoising priors (our proposed prior–★, Noise2Self-UNET, and Noise2Self-FFDNet), pre-trained deep denoising priors (Pretrained-UNet and Pretrained-FFDNet), and classical denoising priors (TV).

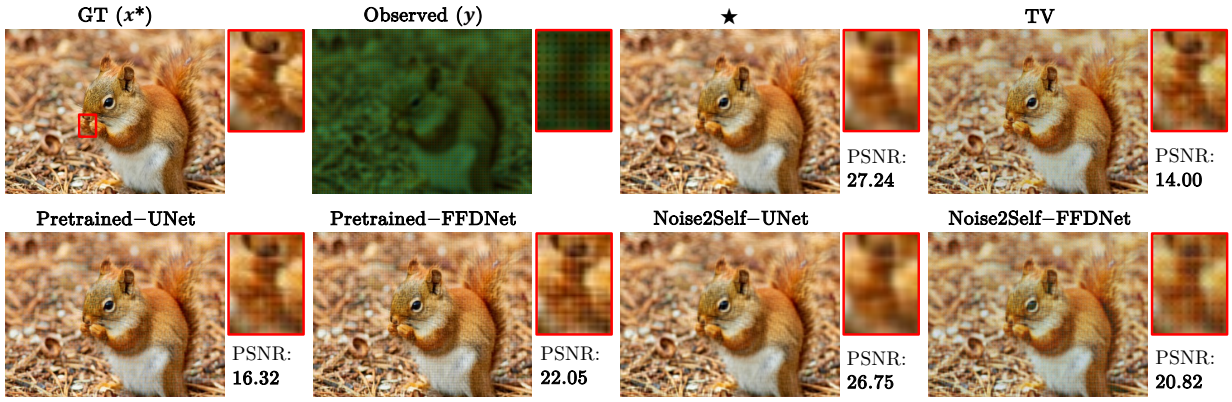


Figure 6: The visualisation comparison of joint deconvolution and demosaicing task on "Squirrel" example between Single-Shot deep denoising priors (our proposed prior–★, Noise2Self-UNET and Noise2Self-FFDNet), pre-trained deep denoising priors (Pretrained-UNet and Pretrained-FFDNet), and classical denoising prior (TV) for denoising in Plug-and-Play algorithm.

In Figure 4, we illustrate a qualitative comparison of our deconvolution algorithm on a 'Giraffe' image against four leading Single-Shot deep denoising priors. The visual fidelity of our method is apparent, particularly in its capacity to reconstruct intricate details, such as the giraffe's fur that reflected in the zoomed views. This attention to detail extends to the preservation of edge sharpness and the subtle gradations, which contribute to a more natural and cohesive image composition. Contrastingly, other methods exhibit varying degrees of blurring and artifact introduction. This qualitative advancement underscores our algorithm's potential to set a remarkable benchmark for image deconvolution.

**Joint Demosaicing and Deconvolution.** Demosaicing is an algorithmic process that reconstructs a full-colour image from the incomplete colour samples output by an image sensor overlaid with a colour filter array (CFA).

The joint demosaicing and deconvolutional process can be represented as: $A(x) = k \otimes (M \odot x)$, where $x$ is the full-colour image, $M$ is the colour filter array (CFA) , $\odot$ denotes element-wise multiplication. The setting of $k$ was same with kernel in deconvolution task.

Based on performances from previous tasks, we note that FFDNet and UNet produce good results. In the more challenging tasks as it is joint demosaicing and deconvolution, these methods, along with classical

Table 3: The performance (PSNR(dB), SSIM) comparison on joint deconvolution and demosaicing of Single-Shot deep denoising priors (★, Noise2Self-UNET, and Noise2Self-FFDNet), pre-trained deep denoising priors (Pretrained-UNet and Pretrained-FFDNet), and classical denoising priors (TV) in Plug-and-Play framework. ★ denotes our proposed denoising prior in the SS-PnP strategy.

| | | Wolf | | Beach | | Mushroom | | Bird | | Head | | Monarch | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| Classic | TV | 17.27 | 0.45 | 16.69 | 0.48 | 17.83 | 0.53 | 18.72 | 0.63 | 17.53 | 0.58 | 16.35 | 0.41 |
| Pretrain | UNet | 17.54 | 0.36 | 18.02 | 0.48 | 18.09 | 0.49 | 19.54 | 0.65 | 19.12 | 0.57 | 17.10 | 0.41 |
| | FFDNet | 25.66 | 0.81 | 24.06 | 0.75 | 25.02 | 0.82 | 24.30 | 0.82 | 24.65 | 0.77 | 21.82 | 0.63 |
| SS-PnP | UNet | 30.54 | 0.90 | 25.52 | 0.77 | 26.86 | 0.87 | 27.75 | 0.87 | 27.96 | 0.81 | 26.31 | **0.89** |
| | FFDNet | 18.45 | 0.28 | 16.17 | 0.25 | 18.78 | 0.34 | 18.15 | 0.45 | 18.68 | 0.29 | 15.39 | 0.26 |
| | ★ | **30.90** | **0.91** | **27.37** | **0.85** | **27.00** | **0.88** | **28.20** | **0.89** | **28.02** | **0.82** | **26.48** | **0.89** |

Total Variation (TV) Jordan (1881), are further compared under data-driven pre-trained and Noise2Self pre-training conditions against our proposed method. Table 3 illustrates that our method outperforms all methods across all categories, achieving the highest PSNR and SSIM scores in most cases. Notably, in the 'Wolf' and 'Monarch' categories, our method significantly leads, reflecting a substantial improvement in both reconstruction accuracy and image quality, as indicated by the PSNR and SSIM metrics respectively. This demonstrates the effectiveness of our approach in handling complex image restoration tasks.

In Figure 5 and 6, our method outperforms other established algorithms in joint deconvolution and demosaicing, particularly in mitigating chromatic aberrations such as red grids or spots. These artifacts, which degrade image quality, are significantly reduced in our approach, leading to a visually coherent result that aligns more closely with the ground truth. Competing methods, including pre-trained networks and classical denoising techniques, often introduce or inadequately suppress such distortions, resulting in inferior colour and contrasts outcomes. Our method advances the visual quality of image restoration, effectively preserving the natural colour and detail fidelity.

Table 4: The performance (PSNR(dB)) comparison of the implicit neural priors on deconvolution (Deconv), super resolution (SR) with 2× and 4× upscalings, and joint deconvolution and demosaicing (Joint) tasks. The ★ denotes our proposed implicit neural prior, which we compared with the SOTA implicit neural representation, SIREN.

| Example | INR Prior | **Deconv** | **SR (2×)** | **SR (4×)** | **Joint** |
|---|---|---|---|---|---|
| **Wolf** | SIREN | 33.05 | 26.08 | 22.26 | 30.83 |
| | ★ | **33.07** | **27.21** | **23.30** | **30.90** |
| **Bird** | SIREN | 30.71 | 25.72 | 21.83 | 28.13 |
| | ★ | **31.07** | **26.44** | **22.06** | **28.20** |
| **Head** | SIREN | 29.82 | 23.69 | 19.57 | 27.79 |
| | ★ | **29.92** | **24.24** | **20.61** | **28.02** |

### 4.3 Ablation study

We provide a further empirical study for comparing our proposed implicit neural network (INR) with the existing classical INR, SIREN Strümpler et al. (2022).

We follow the same experimental settings as described in 4.1, and measure the performance based on Peak Signal-to-Noise Ratio (PSNR). In Table 4, our proposed INR outperforms SIREN in all the 4 tasks: deconvolution, super resolution 2× and 4×, and joint deconvolution and demosaicing. This performance is significant in the super resolution with an average improvement of 0.8dB. The spatial compactness brought by our INR guarantees the representativeness of different levels of feature that pushes the performance of super resolution tasks.

We can observe from Figure 7 that our INR excels in reconstructing the image with smoother edge features and enhanced colour representation. The performance over iteration curves demonstrate our leading advances over the optimisation iterations, meanwhile show the early iteration steps is effective for achieving the best performance.

## 5 Conclusion

Our work pioneers the Single-Shot Plug-and-Play (SS-PnP) method, transforming the use of PnP priors in inverse problem-solving by reducing reliance on large-scale pre-training of denoisers, and using a single instance. We also propose Single-Shot proximal denoisers via implicit neural priors allowing for superior approximation quality for solving inverse problems using only a single instance. We propose a novel function for implicit neural priors that has desirable theoretical guarantees. We also showed that our work generalises well across different tasks, showing empirical stability for single and multiple operators. We then open the door to a new research line for Single-Shot PnP. Whilst this work uses mainly PnP-ADMM due to the well-know properties, and performance in comparison with other algorithms. Future work will include to evaluate our new research line on Single-Shot PnP on different algorithms including but not limited to FISTA Beck & Teboulle (2009), HQS Geman & Yang (1995), and Primal-dual Dantzig et al. (1956) algorithms.

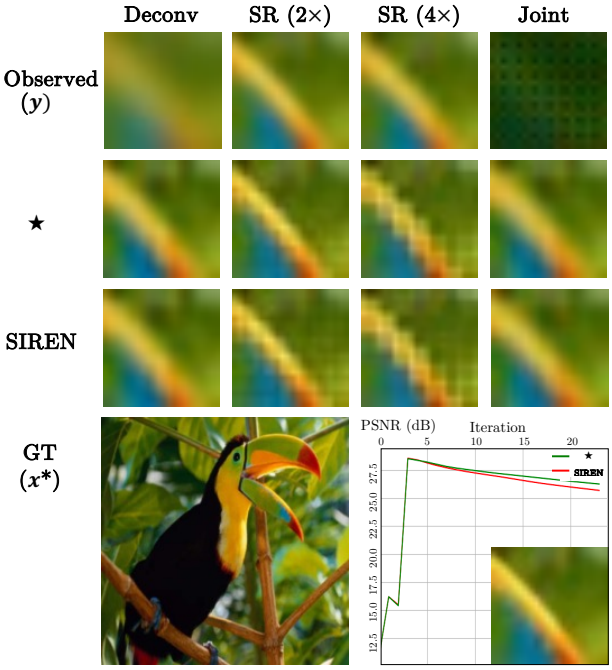


Figure 7: The visualisation comparison of the implicit neural priors (Ours–★ and SIREN) on deconvolution (Deconv), super resolution(SR) with 2× and 4× upscalings, and joint deconvolution and demosaicing (Joint) tasks on "Bird" example. The plot indicates the performance (PSNR(dB)) with both implicit neural priors changing over the ADMM iteration on the joint deconvolution and demosaicing task.

## References

Rizwan Ahmad, Charles A Bouman, Gregery T Buzzard, Stanley Chan, Sizhuo Liu, Edward T Reehorst, and Philip Schniter. Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery. *IEEE signal processing magazine*, 37(1):105–116, 2020.

Ahmet Oğuz Akyüz et al. Deep joint deinterlacing and denoising for single shot dual-iso hdr reconstruction. *IEEE Transactions on Image Processing*, 29:7511–7524, 2020.

Saeed Anwar and Nick Barnes. Densely residual laplacian super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(3):1192–1204, 2020.

Simon Arridge, Peter Maass, Ozan Öktem, and Carola-Bibiane Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019.

Benjamin Attal, Eliot Laidlaw, Aaron Gokaslan, Changil Kim, Christian Richardt, James Tompkin, and Matthew O'Toole. Törf: Time-of-flight radiance fields for dynamic scene view synthesis. *Advances in neural information processing systems*, 34:26289–26301, 2021.

Joshua Batson and Loic Royer. Noise2self: Blind denoising by self-supervision. In *International Conference on Machine Learning*, pp. 524–533. PMLR, 2019.

A Beck and M Teboulle. A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring. In *IEEE*, 2009.

Mario Bertero, Patrizia Boccacci, and Christine De Mol. *Introduction to inverse problems in imaging.* 2021.

Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. 2012.

Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 2, pp. 60–65. Ieee, 2005.

Stanley H Chan, Xiran Wang, and Omar A Elgendy. Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, 2016.

Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8628–8638, 2021.

Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

G. B. Dantzig, Jr Ford, L. R., and D. R. Fulkerson. A primal–dual algorithm. *A Primal Dual Algorithm*, 1956.

Anthony J Devaney. *Mathematical foundations of imaging, tomography and wavefield inversion.* Cambridge University Press, 2012.

Geman and Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 1995.

Samuel Hurault, Arthur Leclaire, and Nicolas Papadakis. Proximal denoiser for convergent plug-and-play optimization with nonconvex regularization. In *International Conference on Machine Learning*, pp. 9483–9505. PMLR, 2022.

Camille Jordan. Sur la series de fourier. *CR Acad. Sci., Paris*, 92:228–230, 1881.

Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2void-learning denoising from single noisy images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2129–2137, 2019.

Zeqiang Lai, Kaixuan Wei, Ying Fu, Philipp Härtel, and Felix Heide. -prox: Differentiable proximal algorithm modeling for large-scale optimization. *ACM Transactions on Graphics (TOG)*, 42(4):1–19, 2023.

Samuli Laine, Tero Karras, Jaakko Lehtinen, and Timo Aila. High-quality self-supervised deep image denoising. *Advances in Neural Information Processing Systems*, 32, 2019.

Victor Lempitsky, Andrea Vedaldi, and Dmitry Ulyanov. Deep image prior. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9446–9454. IEEE, 2018.

Michael T McCann, Kyong Hwan Jin, and Michael Unser. Convolutional neural networks for inverse problems in imaging: A review. *IEEE Signal Processing Magazine*, 34(6):85–95, 2017.

Tim Meinhardt, Michael Moller, Caner Hazirbas, and Daniel Cremers. Learning proximal operators: Using denoising networks for regularizing inverse imaging problems. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1781–1790, 2017.

Nicholas Metropolis and Stanislaw Ulam. The monte carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949.

Shunsuke Ono. Primal-dual plug-and-play image restoration. *IEEE Signal Processing Letters*, 24(8):1108–1112, 2017.

Yuhui Quan, Mingqin Chen, Tongyao Pang, and Hui Ji. Self2self with dropout: Learning self-supervised denoising from single image. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1890–1898, 2020.

Yaniv Romano, Michael Elad, and Peyman Milanfar. The little engine that could: Regularization by denoising (red). *SIAM Journal on Imaging Sciences*, 10(4):1804–1844, 2017.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pp. 234–241. Springer, 2015.

Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, and Wotao Yin. Plug-and-play methods provably converge with properly trained denoisers. In *International Conference on Machine Learning*, pp. 5546–5557. PMLR, 2019.

Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G Baraniuk. Wire: Wavelet implicit neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18507–18516, 2023.

Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems*, 33: 7462–7473, 2020.

Yannick Strümpler, Janis Postels, Ren Yang, Luc Van Gool, and Federico Tombari. Implicit neural representations for image compression. In *European Conference on Computer Vision*, pp. 74–91. Springer, 2022.

Yu Sun, Brendt Wohlberg, and Ulugbek S Kamilov. An online plug-and-play algorithm for regularized image reconstruction. *IEEE Transactions on Computational Imaging*, 5(3):395–408, 2019.

Yu Sun, Jiaming Liu, Mingyang Xie, Brendt Wohlberg, and Ulugbek S Kamilov. Coil: Coordinate-based internal learning for imaging inverse problems. *arXiv preprint arXiv:2102.05181*, 2021.

Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems*, 33:7537–7547, 2020.

Afonso M Teodoro, José M Bioucas-Dias, and Mário AT Figueiredo. A convergent image fusion algorithm using scene-adapted gaussian-mixture-based denoising. *IEEE Transactions on Image Processing*, 28(1): 451–463, 2018.

Singanallur V Venkatakrishnan, Charles A Bouman, and Brendt Wohlberg. Plug-and-play priors for model based reconstruction. In *2013 IEEE global conference on signal and information processing*, pp. 945–948. IEEE, 2013.

Curtis R Vogel. *Computational methods for inverse problems*. SIAM, 2002.

Jiachuan Wang, Shimin Di, Lei Chen, and Charles Wang Wai Ng. Noise2info: Noisy image to information of noise for self-supervised image denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16034–16043, 2023.

Yuehao Wang, Yonghao Long, Siu Hin Fan, and Qi Dou. Neural rendering for stereo 3d reconstruction of deformable tissues in robotic surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 431–441. Springer, 2022.

Kaixuan Wei, Angelica Aviles-Rivero, Jingwei Liang, Ying Fu, Carola-Bibiane Schönlieb, and Hua Huang. Tuning-free plug-and-play proximal algorithm for inverse imaging problems. In *International Conference on Machine Learning*, pp. 10158–10169. PMLR, 2020.

Kaixuan Wei, Angelica Aviles-Rivero, Jingwei Liang, Ying Fu, Hua Huang, and Carola-Bibiane Schönlieb. Tfpnp: Tuning-free plug-and-play proximal algorithms with applications to inverse imaging problems. *The Journal of Machine Learning Research*, 23(1):699–746, 2022.

Yaochen Xie, Zhengyang Wang, and Shuiwang Ji. Noise2same: Optimizing a self-supervised bound for image denoising. *Advances in neural information processing systems*, 33:20320–20330, 2020.

Xin Yuan, Yang Liu, Jinli Suo, and Qionghai Dai. Plug-and-play algorithms for large-scale snapshot compressive imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1447–1457, 2020.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Cycleisp: Real image restoration via improved data synthesis. In *IEEE/CVF conference on computer vision and pattern recognition*, pp. 2696–2705, 2020.

Roman Zeyde, Michael Elad, and Matan Protter. On single image scale-up using sparse-representations. In *Curves and Surfaces: 7th International Conference, Avignon, France, June 24-30, 2010, Revised Selected Papers 7*, pp. 711–730. Springer, 2012.

Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26(7):3142–3155, 2017a.

Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep cnn denoiser prior for image restoration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3929–3938, 2017b.

Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, 27(9):4608–4622, 2018.

Kai Zhang, Wangmeng Zuo, and Lei Zhang. Deep plug-and-play super-resolution for arbitrary blur kernels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1671–1681, 2019.

Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44 (10):6360–6376, 2021.