Training deep learning based image denoisers from undersampled measurements without ground truth and without image prior

Magauiya Zhussip,   Shakarim Soltanayev,   Se Young Chun

Ulsan National Institute of Science and Technology (UNIST), Republic of Korea

{mzhussip,shakarim,sychun}@unist.ac.kr

Abstract

Compressive sensing is a method to recover the original image from undersampled measurements. In order to overcome the ill-posedness of this inverse problem, image priors are used such as sparsity, minimal total-variation, or self-similarity of images. Recently, deep learning based compressive image recovery methods have been proposed and have yielded state-of-the-art performances. They used data-driven approaches instead of hand-crafted image priors to regularize ill-posed inverse problems with undersampled data. Ironically, training deep neural networks (DNNs) for them requires "clean" ground truth images, but obtaining the best quality images from undersampled data requires well-trained DNNs. To resolve this dilemma, we propose novel methods based on two well-grounded theories: denoiser-approximate message passing (D-AMP) and Stein's unbiased risk estimator (SURE). Our proposed methods were able to train deep learning based image denoisers from undersampled measurements without ground truth images and without additional image priors, and to recover images with state-of-the-art qualities from undersampled data. We evaluated our methods for various compressive sensing recovery problems with Gaussian random, coded diffraction pattern, and compressive sensing MRI measurement matrices. Our proposed methods yielded state-of-the-art performances for all cases without ground truth images. Our methods also yielded comparable performances to the methods with ground truth data.

1. Introduction

Compressive sensing (CS) has provided ways to sample and to compress signals at the same time with relatively long signal reconstruction time [10, 15]. The idea of combining signal acquisition and compression immediately drew great attention in the application areas such as MRI [29, 38], CT [11], hyperspectral imaging [49, 50], coded aperture imaging [2], radar imaging [36] and radio

astronomy [41]. CS applications have been investigated extensively for the last decade and now some systems are commercialized for practical usages such as low-dose CT and accelerated MR.

CS is modeled as a linear equation for the measurement:

$$y = Ax + \epsilon \tag{1}$$

where $y \in \mathbb{R}^M$ is a measurement vector, $A \in \mathbb{R}^{M \times N}$ is a sensing matrix with $M \ll N$, $x \in \mathbb{R}^N$ is an unknown signal to reconstruct, and $\epsilon \in \mathbb{R}^M$ is a noise vector. It is a challenging ill-posed inverse problem to estimate x from the undersampled measurements y with $M \ll N$.

Sparsity has been investigated as prior to regularize the ill-posed problem of CS recovery. CS theories allow to use l_1 norm for good signal recovery instead of l_0 norm [10, 15]. Minimizing l_1 norm is advantageous for large-scale inverse problems since l_1 norm is convex so that conventional convex optimization can be used. There have been many convex optimization algorithms for solving CS recovery problems with non-differentiable l_1 norm such as iterative shrinkage-thresholding algorithm (ISTA), fast iterative
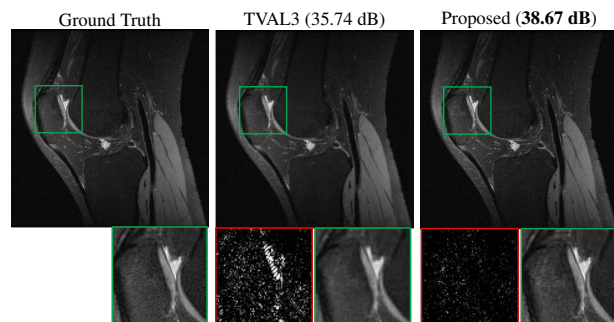

Figure 1: Ground truth MR image from fully-sampled data (left), reconstructed MR images from 50%-sampled data using conventional TV image prior (middle, [29]) and our proposed deep learning based method without ground truth (right). Our proposed method yielded significantly better result than conventional method and result comparable to the ground truth with very small residual error (red box).

shrinkage-thresholding algorithm (FISTA) [3], alternating direction minimization (ADM) [44], and approximate message passing (AMP) [16], to name a few.

The signal x itself is not usually sparse, but a transformed signal is often sparse. Signals/images are sparse in the wavelet domain and/or discrete cosine transform (DCT) domain. In high-resolution imaging, images have sparse edges that are often promoted by minimizing total variation (TV) [28]. Sparse MR image recovery used both wavelet and TV priors [29] or dictionary learning prior from highly undersampled measurements [38]. Similarly, CS color video and depth recovery used both wavelet and DCT [46]. Hyperspectral imaging utilized manifold-structured sparsity prior [49] or reweighted Laplace prior [50]. Self-similarity is also used for CS image recovery such as NLR-CS [14] and denoiser-AMP (D-AMP) [32]. D-AMP utilized powerful modern denoisers such as BM3D [12] and has recently been extended to sparse MRI [17].

1.1. Deep learning in compressive image recovery

Deep learning with massive amount of data has revolutionized many computer vision tasks [26]. It has also influenced many low level computer vision tasks such as image denoising [8, 22, 40, 42, 48, 35] and CS recovery [20, 21, 23, 25, 33, 43, 47]. There are largely two different approaches using deep neural networks (DNNs) for CS recovery. One is to use a deep network to directly map from an initially recovered low-quality image from compressive samples to a high-quality ground truth image [23, 25]. The other approach for deep learning based CS image recovery is to use DNN structures that unfold optimization algorithms and learned image priors, inspired by learned ISTA (LISTA) [19]. In sparse MRI, ADMM-Net [45] and variational network [21] were proposed with excellent performances. Both methods learned parametrized shrinkage functions as well as transformation operators for sparse representation from training data. Recently, instead of using explicit parametrization in shrinkage operator, DNNs were used to unfold optimizations for CS recovery such as learned D-AMP (LDAMP) [33], ISTA-Net [47], CNN-projected gradient descent for CT [20], and Laplacian pyramid reconstructive adversarial network [43]. Utilizing generative adversarial network (GAN) for CS was also investigated [4]. These methods have one important requirement: "clean" ground truth images must be available for training.

1.2. Deep learning without ground truth

Most deep learning based data-driven approaches for CS image recovery solve the ill-posed inverse problem with undersampled data by using DNNs. Ironically, training DNNs for them requires "clean" ground truth images, but obtaining the best quality images from undersampled data requires well-trained DNNs. It is often expensive or infeasible to acquire clean data, for example, in medical imaging (long acquisition for MR, high radiation dose for CT) or hyperspectral imaging. Here, we address this dilemma.

Recently, there have been a few attempts to train DNNs for low-level computer vision tasks in unsupervised ways. Lehtinen *et al.* proposed noise2noise to train DNNs for image denoising, inpainting, and MR reconstruction [27]. This work implemented MR reconstruction using a direct mapping instead of unfolding optimization scheme. However, this was not evaluated with various CS applications and it requires two contaminated data for each image, which may not be available in some cases. Bora *et al.* proposed AmbientGAN, a training method for GAN with contaminated images and applied it to CS image recovery [4, 5]. However, AmbientGAN was trained with artificially contaminated images, rather than with CS measurements. Moreover, the method of [4] is limited to $i.i.d.$ Gaussian measurement matrix theoretically and was evaluated with relatively low-resolution images. Soltanayev *et al.* proposed a Stein's unbiased risk estimator (SURE) based training method for deep learning based denoisers [39]. This method requires only one realization, but it is limited to $i.i.d.$ Gaussian noise. Moreover, it is not straightforward to extend this work to CS image recovery with measurements.

We propose unsupervised training methods for deep learning based CS image recovery based on two well-grounded theories: D-AMP and SURE. Our proposed methods were able to train DNN based image denoisers from undersampled measurements without ground truth images, and to recover images with state-of-the-art qualities from undersampled data. Here are the contributions of this work:

1) Proposing a method to train DNN denoisers from undersampled measurements without ground truth and without additional image priors. Only one realization for each measurement was required. An accurate noise estimation method was also developed for training deep denoisers.

2) Proposing a CS image recovery method by modifying LDAMP to have up to1 denoiser instead of 9 denoisers with comparable performance to reduce training time.

3) Extensive evaluations of the proposed method using high-resolution natural images and MR images for the CS recovery problems with Gaussian random, coded diffraction pattern, and realistic CS MR measurement matrices.

2. Background

2.1. Denoiser-based AMP (D-AMP)

D-AMP is an algorithm designed to solve CS problems where one needs to recover image vector x from the set of measurements y using prior information about x. Based on the model (1), the problem can be formulated as:

$$\min_{x} \|y - Ax\|_2^2 \quad \text{subject to} \quad x \in C \qquad (2)$$

Algorithm 1: (Learned) D-AMP algorithm [32, 33]

> **input** : $\boldsymbol{x}_0 = \boldsymbol{0}, \boldsymbol{y}, \boldsymbol{A}$
> **1 for** $t = 0$ *to* T **do**
> **2** $\boldsymbol{b}_t \leftarrow \boldsymbol{z}_{t-1} \mathrm{div} \boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_{t-1})}(\boldsymbol{x}_{t-1} + \boldsymbol{A}^H \boldsymbol{z}_{t-1})/M$
> **3** $\boldsymbol{z}_t \leftarrow \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{b}_t$
> **4** $\hat{\sigma}_t \leftarrow \|\boldsymbol{z}_t\|_2 / \sqrt{M}$
> **5** $\boldsymbol{x}_{t+1} \leftarrow \boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\boldsymbol{x}_t + \boldsymbol{A}^H \boldsymbol{z}_t)$
> **6 end**
> **output:** $\boldsymbol{x}_T$

where C is the set of natural images. D-AMP solves (2) relying on approximate message passing (AMP) theory. It employs appropriate Onsager correction term $\boldsymbol{b}_t$ at each iteration, so that $\boldsymbol{x}_t + \boldsymbol{A}^H \boldsymbol{z}_t$ in Algorithm 1 becomes close to the ground truth image plus $i.i.d.$ Gaussian noise. D-AMP can utilize any denoiser as a mapping operator $\boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\cdot)$ in CS recovery (Algorithm 1) for reducing $i.i.d.$ Gaussian noise as far as the divergence of denoiser can be obtained.

D-AMP [31] first utilized conventional state-of-the-art denoisers such as BM3D [12] for $\boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\cdot)$ in Algorithm 1. Given a standard deviation of noise $\hat{\sigma}_t$ at iteration t, BM3D was applied to a noisy image $\boldsymbol{x}_t + \boldsymbol{A}^H \boldsymbol{z}_t$ to yield estimated image $\boldsymbol{x}_{t+1}$. Since BM3D can not be represented as a linear function, analytical form for divergence of this denoiser is not available for obtaining Onsager term. This issue was resolved by using Monte-Carlo (MC) approximation for divergence term $\mathrm{div}\boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\cdot)$ [37]: For $\epsilon > 0$ and $\tilde{\boldsymbol{n}}$ is a standard normal random vector,

$$\mathrm{div}\boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\cdot) \approx \frac{\tilde{\boldsymbol{n}}'}{\epsilon}\left(\boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\cdot + \epsilon\tilde{\boldsymbol{n}}) - \boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\cdot)\right). \quad (3)$$

Recently, LDAMP was proposed to use deep learning based denoisers for $\boldsymbol{D}_{\boldsymbol{w}(\hat{\sigma}_t)}(\cdot)$ in Algorithm 1. Nine DNN denoisers were trained with noiseless ground truth data for different noise levels. LDAMP consists of 10 D-AMP layers (iterations) and DnCNN [48] was used in each layer as a denoiser operator. Unlike other unrolled neural network versions of iterative algorithms such as learned-AMP [6] and LISTA [19], LDAMP exploited imaging system models, which are fixed $\boldsymbol{A}$ and $\boldsymbol{A}^H$ operators while the parameters for DnCNN denoisers were trained with ground truth data in image domain.

2.2. Stein's unbiased risk estimator (SURE) based deep neural network denoisers

Over the past years, DNN based denoisers have been well investigated [8, 22, 40, 42, 48, 35] and often outperformed conventional denoisers such as BM3D [12] and nonlocal filtering [7, 34]. DNN denoisers such as DnCNN [48] yielded state-of-the-art denoising performance at multiple noise levels and are typically trained by minimizing the mean square error (MSE) between the output image of denoiser and the noiseless ground truth image:

$$\frac{1}{K}\sum_{j=1}^{K}\|\boldsymbol{D}_{\boldsymbol{w}(\sigma)}(\boldsymbol{z}^{(j)}) - \boldsymbol{x}^{(j)}\|^2 \quad (4)$$

where $\boldsymbol{z} \in \mathbb{R}^N$ is a noisy image of the ground truth image $\boldsymbol{x}$ contaminated with $i.i.d.$ Gaussian noise with zero mean and fixed σ^2 variance, $\boldsymbol{D}_{\boldsymbol{w}(\sigma)}(\cdot)$ is a deep learning based denoiser with large-scale parameters $\boldsymbol{w}$ to train, and $(\boldsymbol{z}^{(1)}, \boldsymbol{x}^{(1)}), \dots, (\boldsymbol{z}^{(K)}, \boldsymbol{x}^{(K)})$ is a training dataset with K samples in image domain.

Recently, a method to train deep learning based denoisers only with noisy images was proposed [39]. Instead of minimizing MSE, the following Monte-Carlo Stein's unbiased risk estimator (MC-SURE) that approximates MSE was minimized with respect to large-scale weights $\boldsymbol{w}$ in the DNN without noiseless ground truth images:

$$\frac{1}{K}\sum_{j=1}^{K}\|\boldsymbol{z}^{(j)} - \boldsymbol{D}_{\boldsymbol{w}(\sigma)}(\boldsymbol{z}^{(j)})\|^2 - N\sigma^2 + \\ \frac{2\sigma^2\tilde{\boldsymbol{n}}'}{\epsilon}\left(\boldsymbol{D}_{\boldsymbol{w}(\sigma)}(\boldsymbol{z}^{(j)} + \epsilon\tilde{\boldsymbol{n}}) - \boldsymbol{D}_{\boldsymbol{w}(\sigma)}(\boldsymbol{z}^{(j)})\right). \quad (5)$$

In compressive image recovery applications, there are often cases where no ground truth data or no Gaussian contaminated images are available, but only compressive samples in measurement domain are available for training. However, it is not straightforward to use MSE or MC-SURE based deep denoiser networks for CS image recovery unless additional image priors are used. The goal of this article is to propose a method to train DNN denoisers directly from compressive samples without additional image prior and to simultaneously recover images.

3. Method

3.1. Training deep denoisers from undersampled measurements without ground truth

Our proposed method exploits D-AMP (LDAMP) [32, 33] to yield Gaussian noise contaminated images during compressive image recovery from large-scale undersampled measurements and train a single DNN denoiser with these noisy images at different noise levels using MC-SURE based denoiser learning [39]. Since Onsager correction term in D-AMP allows $\boldsymbol{x} + \boldsymbol{A}^H \boldsymbol{z}$ term to be close to the ground truth image plus Gaussian noise, we conjecture that these can be utilized for MC-SURE based denoiser training. We further investigated this in the next subsection. Our joint algorithm is detailed in Algorithm 2. Note that for large-scale compressive measurements $\boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(K)}$, both images $\hat{\boldsymbol{x}}_L^{(1)}, \dots, \hat{\boldsymbol{x}}_L^{(K)}$ and trained denoising deep network $\boldsymbol{D}_{\boldsymbol{w}_L(\sigma)}(\cdot)$ were able to be obtained. After training,

fast and high performance compressive image recovery was possible without further training of deep denoising network.

The original LDAMP [33] utilized 9 DnCNN denoisers trained on "clean" images for different noise levels ($\sigma = 0 \sim 10, 10 \sim 20, 20 \sim 40, 40 \sim 60, 60 \sim 80, 80 \sim 100, 100 \sim 150, 150 \sim 300, 300 \sim 500$). However, we found that training a single DnCNN denoiser is enough to achieve almost the same results (see Table 2 in the supplementary material). The DnCNN network was pre-trained with reconstructed images using D-AMP with BM3D plus Gaussian noise with $\sigma \in [0, 55]$. The pre-trained DnCNN blind denoiser $\boldsymbol{D}_{\boldsymbol{w}_l(\hat{\sigma}_t)}$ cleans $\boldsymbol{x}_{t-1} + \boldsymbol{A}^H \boldsymbol{z}_{t-1}$ with noise level between $[0, 55]$ (line 8 in Algorithm 2), while $BM3D_{\hat{\sigma}_t}$ is utilized for higher level noise reduction (line 10 in Algorithm 2). Depending on a sampling ratio and forward operator $\boldsymbol{A}$, initial 2-4 iterations are usually required for BM3D to decrease the noise level sufficient enough for DnCNN. Then, after T iterations, the set of training data $\boldsymbol{s}_l^{(1)}, \ldots, \boldsymbol{s}_l^{(K)}$ can be generated using LDAMP with pre-trained deep denoiser. Those noisy training images were used for further training the DnCNN with MC-SURE.

Worth to note that the noise level range for DnCNN is subject to change depending on a particular problem. For example, we found that for $i.i.d.$ Gaussian and CDP matrices, training DnCNN with $\sigma \in [0, 55]$ is optimal, while for CS MRI case, the range is shortened to $\sigma \in [0, 10]$.

3.2. Accuracy of standard deviation estimation for MC-SURE based denoiser learning

In D-AMP and LDAMP [32, 33], noise level was estimated in measurement domain using

$$\hat{\sigma}_t \leftarrow \|\boldsymbol{z}_t\|_2 / \sqrt{M}. \tag{6}$$

The accuracy of this estimation was not critical for D-AMP or LDAMP since denoisers in both methods were not sensitive to different noise levels. However, accurate noise level estimation was quite important for MC-SURE based deep denoiser network learning. We investigated the accuracy of (6). It turned out that the accuracy of noise level estimation depends on measurement matrices.

With $i.i.d.$ Gaussian measurement matrix $\boldsymbol{A}$, (6) was very accurate and comparable to the ground truth standard deviation that was obtained from the true residual $(\boldsymbol{x}_t + \boldsymbol{A}^H \boldsymbol{z}_t) - \boldsymbol{x}_{true}$. However, with coded diffraction pattern measurement matrix $\boldsymbol{A}$ that yields complex measurements, it turned out that (6) yielded over-estimated noise level for multiple examples. Since the image $\boldsymbol{x}_t$ is real, we propose a new standard estimation method for D-AMP:

$$\hat{\sigma}_t \leftarrow \| \operatorname{Re}(\boldsymbol{A}^H \boldsymbol{z}_t)\|_2 / \sqrt{N}. \tag{7}$$

We performed comparison studies between (6), (7), and the ground truth from true residual $(\boldsymbol{x}_t + \boldsymbol{A}^H \boldsymbol{z}_t) - \boldsymbol{x}_{true}$

Algorithm 2: Simultaneous LDAMP and MC-SURE deep denoiser learning algorithm

input : $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(K)}, \boldsymbol{A}$

1   **for** $l = 1$ *to* L **do**
2   **for** $k = 1$ *to* K **do**
3   **for** $t = 0$ *to* T **do**
4      $\boldsymbol{b}_t \leftarrow$
      $\boldsymbol{z}_{t-1}\mathrm{div}\boldsymbol{D}_{\boldsymbol{w}_l(\hat{\sigma}_{t-1})}(\boldsymbol{x}_{t-1} + \boldsymbol{A}^H \boldsymbol{z}_{t-1})/M$
5      $\boldsymbol{z}_t \leftarrow \boldsymbol{y}^{(k)} - \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{b}_t$
6      $\hat{\sigma}_t \leftarrow \|\boldsymbol{z}_t\|_2 / \sqrt{M}$
7   **if** $\hat{\sigma}_t \leq 55.$ **then**
8       $\boldsymbol{x}_{t+1} \leftarrow \boldsymbol{D}_{\boldsymbol{w}_l(\hat{\sigma}_t)}(\boldsymbol{x}_t + \boldsymbol{A}^H \boldsymbol{z}_t)$
9   **else**
10       $\boldsymbol{x}_{t+1} \leftarrow BM3D_{\hat{\sigma}_t}(\boldsymbol{x}_t + \boldsymbol{A}^H \boldsymbol{z}_t)$
11   **end**
12   **end**
13     $\hat{\boldsymbol{x}}_l^{(k)} \leftarrow \boldsymbol{x}_{T+1}$
14     $\boldsymbol{s}_l^{(k)} \leftarrow \boldsymbol{x}_T + \boldsymbol{A}^H \boldsymbol{z}_T$
15   **end**
16   Train $\boldsymbol{D}_{\boldsymbol{w}_l(\sigma)}(\cdot)$ with $\boldsymbol{s}_l^{(1)}, \ldots, \boldsymbol{s}_l^{(K)}$ at different noise levels σ
17   **end**

output: $\hat{\boldsymbol{x}}_L^{(1)}, \ldots, \hat{\boldsymbol{x}}_L^{(K)}, \boldsymbol{D}_{\boldsymbol{w}_L(\sigma)}(\cdot)$

---

and found that they are all very similar for $i.i.d.$ Gaussian measurement matrix, but our proposed method (7) yielded more accurate estimates of standard deviation than previous method (6). Figure 2 illustrates the accuracy of our estimator compared to previous one. When normalizing the true residual, using accurate sigma estimation yields good fitting to standard normal density (red line). Normalized histogram of true residual using ground truth and our proposed standard deviation estimation yielded good fitting to that, but previous estimation method yielded sharper histogram, which indicates that previous method overestimates noise level. Our proposed estimation was critical for the high performance of our proposed method with CDP measurement matrix (see Table 1 in the supplementary material).

Moreover, we found out that proposed noise estimator (7) also can be applied to a CS-MRI case, when k-space data is not highly undersampled. Therefore, for a sampling rate of larger than 35-40%, true residual follows a Gaussian noise, which can be accurately measured by (7) and further utilized for training deep denoisers with MC-SURE.

4. Simulation Results

4.1. Setup

Datasets We used images from DIV2K [1], Berkeley's BSD-500 [30] datasets, and standard test images for train-
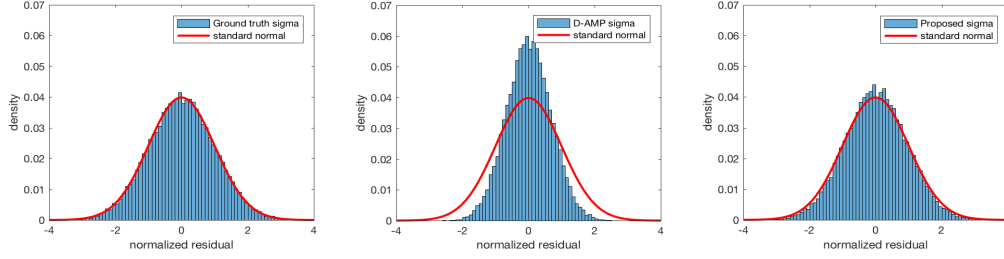
Figure 2: Normalized residual histograms of "Boat" image after 10 iterations using LDAMP-BM3D for CDP matrix. Normalization was done with estimated sigma from (a) true residual (b) $\boldsymbol{z}_T$ (D-AMP) and (c) $\mathrm{Re}(\boldsymbol{A}^H \boldsymbol{z}_T)$ (Proposed).

ing and testing our proposed method on $i.i.d.$ Gaussian and CDP matrices. Training dataset was comprised of all 500 images from BSD-500, while a test set of 100 images included 75 randomly chosen images from DIV2K dataset and 25 standard test images. Since the proposed method used measurement data and fixed linear operator for image reconstruction, all test and train images had to have the same size. Thus, all images were subsampled and cropped to the size of 180×180 and then compressively sampled using the forward model $\boldsymbol{A}$ to generate CS measurement data.

For CS-MRI reconstruction, Stanford dataset with 3D FSE (fast spin echo) [18] was pulled from the open repository at http://mridata.org/. The knee dataset included 20 patients each having 256 slices of 320×320 images. Among 20 cases of knee data, 3 cases were used for training and 1 case for testing. Images were transformed to k-space measurements and then subsampled with realistic radial sampling patterns at various sampling rates.

We implemented all methods on the Tensorflow framework and used Adam optimization [24] with the learning rate of 10^{-3}, which was dropped to 10^{-4} after 40 epochs and further trained for 10 epochs. The batch size was set to 128 and training the DnCNN denoiser took approximately 12-14 hours on one NVIDIA Titan X (Pascal).

Initialization of DnCNN denoiser For given measurement data $\boldsymbol{y}$ from BSD-500 and linear operator $\boldsymbol{A}$, initial images were firstly reconstructed using a conventional CS recovery algorithm, BM3D-AMP. Even though the quality of these initial images were not close to the ground truth images, they still provided good pre-training data for DnCNN denoisers. Recovered images were rescaled, cropped, flipped, and rotated to generate 298,000 image patches whose sizes are 50×50 . These patches were used as a ground truth to pre-train DnCNN denoiser with MSE. Since our approach does not require dataset with ground truth, it is possible to use measurement data from the test set. Thus, we also generated 357,600 50×50 patches from reconstructed test and train images. Our DnCNN denoiser was trained for $\sigma \in [0, 55]$ noise level range with either

training patches only or training and testing patches together. The former pre-trained DnCNN denoiser in the LDAMP framework is denoted by "LDAMP BM3D" and the latter pre-trained DnCNN with LDAMP is denoted by "LDAMP BM3D-T".

BM3D-AMP-MRI was specifically tailored for CS-MRI reconstruction [17] and thus yielded significantly better results than conventional BM3D-AMP. Therefore, k-space knee dataset was reconstructed using it and then the resulted images were rescaled, cropped, flipped, and rotated to generate 267,240 and 350,320 50×50 patches for training LDAMP BM3D and LDAMP BM3D-T, respectively. We trained DnCNN denoisers for $\sigma \in [0, 10]$ noise range.

Training LDAMP SURE Firstly, LDAMP SURE was run $T = 10$ iterations using pre-trained DnCNN denoiser and BM3D. At the last iteration, we collected images and estimated noise standard deviation with (7). Then, all images with noise levels in $[0, 55]$ range (CS-MRI case: $\sigma \in [0, 10]$) were grouped into one set, while images with larger noise levels were replaced by Gaussian noise added BM3D-AMP recovered images. Thus, we have the dataset of all images with $\sigma \in [0, 55]$ (CS-MRI case: $\sigma \in [0, 10]$) to train DnCNN denoiser with MC-SURE. These steps were repeated L times to further improve the performance of our proposed method. Although training DnCNN with MC-SURE involves estimation of a noise standard deviation for an entire image, we assume that a patch from an image has the same noise level as the image itself. Thus, we generated patches without using rescaling to avoid noise distortion to train LDAMP SURE.

To train DnCNN with SURE, we initialized DnCNN with the weights of pre-trained DnCNN and trained it using Adam optimizer [24] with learning rate of 10^{-4} and batch size 128 for 10 epochs. Then, we decreased learning rate to 10^{-5} and trained it for another 10 epochs. Training process took about 3 hours for LDAMP SURE and about 4 hours for LDAMP SURE-T. We empirically found that after $L=2$ iterations (line 1 in Algorithm 2) of training LDAMP SURE, the results converge for both CDP and $i.i.d.$ Gaussian cases,

while for CS-MRI, $L = 1$.

The accuracy of MC-SURE approximation depends on the selection of constant value ϵ, which is directly proportional to σ [13, 39]. Therefore, for training DnCNN with SURE, ϵ value was calculated for each patch based on its noise level (see Section 1 in the supplementary material).

4.2. Results

Gaussian measurement matrix We compared our proposed LDAMP SURE with the state-of-the-art CS methods that do not require ground truth data such as BM3D-AMP[31], NLR-CS[14], and TVAL3[28]. BM3D-AMP was used with default parameters and run for 30 iterations to reduce high variation in the results, although PSNR[1] approached its maximum after 10 iterations [32]. The proposed LDAMP SURE algorithm was run 30 iterations but also showed convergence after 8-10 iterations. NLR-CS was initialized with 8 iterations of BM3D as justified in [32], while TVAL3 was set to its default parameters. Also, we included the results of LDAMP with a single DnCNN (denoted as LDAMP MSE) that was trained on ground truth images to see the performance gap.

From Table 1, proposed LDAMP SURE and LDAMP SURE-T outperformed other methods at higher CS ratios by 0.26-0.46 dB, while at a highly undersampled case, it is inferior to NLR-CS. Nevertheless, it is clear that SURE based LDAMP is able to improve the performance of pre-trained LDAMP BM3D and surpasses BM3D-AMP by 0.28-1.56 dB. In Figure 3, reconstructions of all methods on a test image are represented for 25% sampling ratio. Proposed LDAMP SURE and LDAMP SURE-T provide sharper edges and preserve more details.

In terms of run time, the dominant source of computation comes from using BM3D denoiser at initial iterations, while DnCNN takes less than a second for inference. LDAMP SURE utilizes CPU for BM3D and GPU for DnCNN. Consequently, proposed LDAMP SURE was faster than BM3D-AMP, NLR-CS, and TVAL3 methods.

Coded diffraction pattern measurements LDAMP SURE was tested with randomly sampled coded diffraction pattern [9] and yielded the best quantitative performance at higher sampling rates (see Table 2 and Figure 4). LDAMP SURE and LDAMP SURE-T achieved about 1.8 dB performance gain over BM3D-AMP. However, at extremely low sampling ratio, our method slightly falls behind TVAL3. LDAMP SURE requires better data than BM3D-AMP reconstructed images from highly undersampled data to pretrain DnCNN. Therefore, one way to surpass TVAL3 at the highly undersampled case is to pretrain DnCNN with TVAL3 images (see Table 3 in supplementary material).

---

[1]PSNR stands for peak signal-to-noise ratio and is calculated by following expression: $10log_{10}(\frac{255^2}{mean(\hat{x}-x_{gt})^2})$ for pixel range $\in [0-255]$

CS MR measurement matrix LDAMP SURE was applied to CS MRI reconstruction problem to demonstrate its generality and to show its performance on images that contain structures different from natural image dataset. We compared LDAMP SURE with state-of-the-art BM3D-AMP-MRI [17] for CS-MR image reconstruction without ground truth data along with TVAL3, BM3D-AMP, and dictionary learning based DL-MRI [38]. Average image recovery PSNRs and run times are tabulated in Table 3. Figure 5 shows that our proposed method yielded state-of-the-art performance, close to the ground truth. The results reveal that proposed LDAMP SURE-T outperforms existing algorithms in all sampling ratios.

5. Discussion and Conclusion

We proposed methods for unsupervised training of image denoisers with undersampled CS measurements. Our methods simultaneously performed CS image recovery and DNN denoiser learning. Our proposed method yielded better image quality than conventional methods at higher sampling rates for $i.i.d$ Gaussian, CDP, and CS MR measurements. Thus, it may be possible that this work can be helpful for areas where obtaining ground truth images is challenging such as hyperspectral or medical imaging.

Note that training deep learning based image denoisers from undersampled data still seems to require to contain enough information in the undersampled measurements. Tables 1 and 2 show that only 5% of the full samples was not enough to achieve state-of-the-art performance possibly due to lack of information in the measurement. Note also that since we assume a single CS measurement for each image and evaluated with various CS matrices with high-resolution images, it was not possible to compare our method to noise2noise [27] and AmbientGAN [4, 5]. Lastly, our proposed method can potentially be used with more advanced deep denoisers for potentially better performance as far as they are trainable with MC SURE loss [39].

Acknowledgments

Method	Training Time	$\frac{M}{N} = 5\%$		$\frac{M}{N} = 15\%$		$\frac{M}{N} = 25\%$	
		PSNR	Time	PSNR	Time	PSNR	Time
TVAL3	N/A	20.46	9.71	24.14	22.96	26.77	34.87
NLR-CS	N/A	**21.88**	128.73	27.58	312.92	31.20	452.23
BM3D-AMP	N/A	21.40	25.98	26.74	24.21	30.10	23.08
LDAMP BM3D	10.90 hrs	21.41	8.98	27.54	3.94	31.20	2.89
LDAMP BM3D-T	14.30 hrs	21.42	8.98	27.61	3.94	31.32	2.89
LDAMP SURE	15.05 hrs	21.44	8.98	27.65	3.94	31.46	2.89
LDAMP SURE-T	17.97 hrs	21.68	**8.98**	**27.84**	**3.94**	**31.66**	**2.89**
LDAMP MSE	10.17 hrs	22.07	8.98	27.78	3.94	31.65	2.89

Table 1: Average PSNRs (dB) and run times (sec) of 100 180×180 image reconstructions for i.i.d. Gaussian measurements case (no measurement noise) at various sampling rates ($M/N \times 100\%$).

Method	Training Time	$\frac{M}{N} = 5\%$		$\frac{M}{N} = 15\%$		$\frac{M}{N} = 25\%$	
		PSNR	Time	PSNR	Time	PSNR	Time
TVAL3	N/A	**22.57**	**0.85**	27.99	**0.75**	32.82	**0.67**
NLR-CS	N/A	19.00	93.05	22.98	86.90	31.24	119.70
BM3D-AMP	N/A	21.66	22.15	27.29	22.28	31.40	17.00
LDAMP BM3D	10.56 hrs	21.97	23.43	28.04	7.01	31.65	2.71
LDAMP BM3D-T	12.67 hrs	21.93	23.43	28.01	7.01	32.12	2.71
LDAMP SURE	15.22 hrs	22.18	23.43	29.14	7.01	33.26	2.71
LDAMP SURE-T	17.61 hrs	22.06	23.43	**29.17**	7.01	**33.51**	2.71
LDAMP MSE	10.17 hrs	22.12	23.43	28.87	7.01	33.88	2.71

Table 2: Average PSNRs (dB) and run times (sec) of 100 180x180 image reconstructions for CDP measurements case (no measurement noise) at various sampling rates ($M/N \times 100\%$).

Method	Training Time	$\frac{M}{N} = 40\%$		$\frac{M}{N} = 50\%$		$\frac{M}{N} = 60\%$	
		PSNR	Time	PSNR	Time	PSNR	Time
TVAL3	N/A	36.76	**0.58**	37.13	**0.24**	38.35	**0.21**
DL-MRI	N/A	36.60	98.51	37.81	97.58	39.13	99.44
BM3D-AMP-MRI	N/A	37.42	14.76	38.94	15.00	40.51	15.36
BM3D-AMP	N/A	36.15	96.23	36.29	84.34	39.53	98.01
LDAMP BM3D	9.31 hrs	37.12	6.26	38.63	6.14	39.53	6.10
LDAMP BM3D-T	12.41 hrs	37.65	6.26	38.92	6.14	39.87	6.10
LDAMP SURE	12.04 hrs	37.40	6.26	38.70	6.14	40.62	6.10
LDAMP SURE-T	16.05 hrs	**37.77**	6.26	**39.09**	6.14	**40.71**	6.10

Table 3: Average PSNRs (dB) and run times (sec) of 100 180x180 image reconstructions for CS-MRI measurements case (no measurement noise) at various sampling rates ($M/N \times 100\%$).
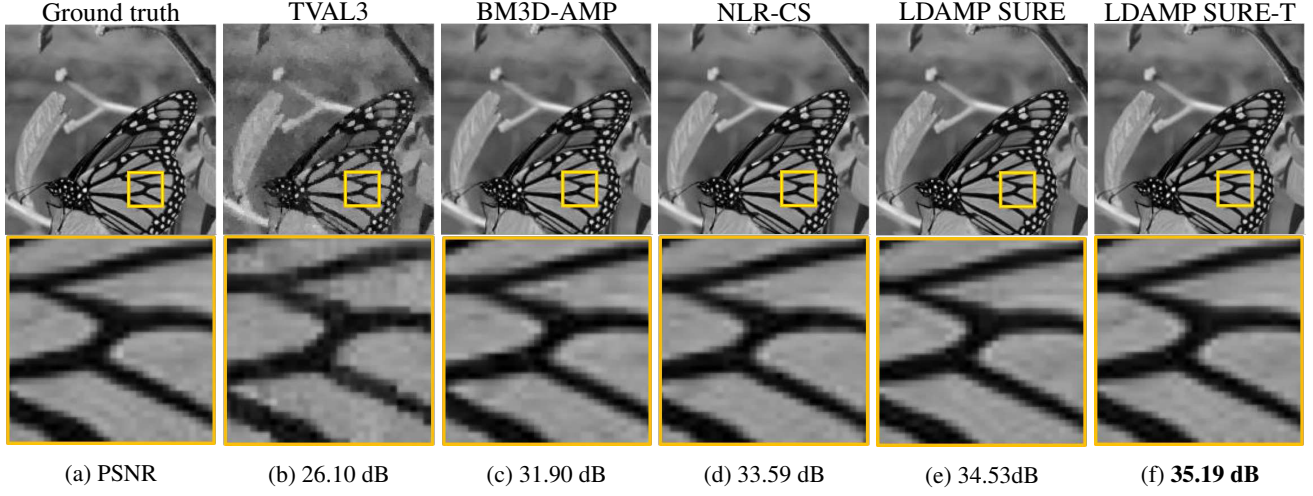
Ground truth	TVAL3	BM3D-AMP	NLR-CS	LDAMP SURE	LDAMP SURE-T
(a) PSNR	(b) 26.10 dB	(c) 31.90 dB	(d) 33.59 dB	(e) 34.53dB	(f) **35.19 dB**

Figure 3: Reconstructions of 180×180 test "Butterfly" image with i.i.d. Gaussian matrix with $M/N = 0.25$ sampling rate.

Ground truth	TVAL3	BM3D-AMP	NLR-CS	LDAMP SURE	LDAMP SURE-T
(a) PSNR	(b) 27.52 dB	(c) 24.08 dB	(d) 22.29 dB	(e) **29.17 dB**	(f) 28.92 dB
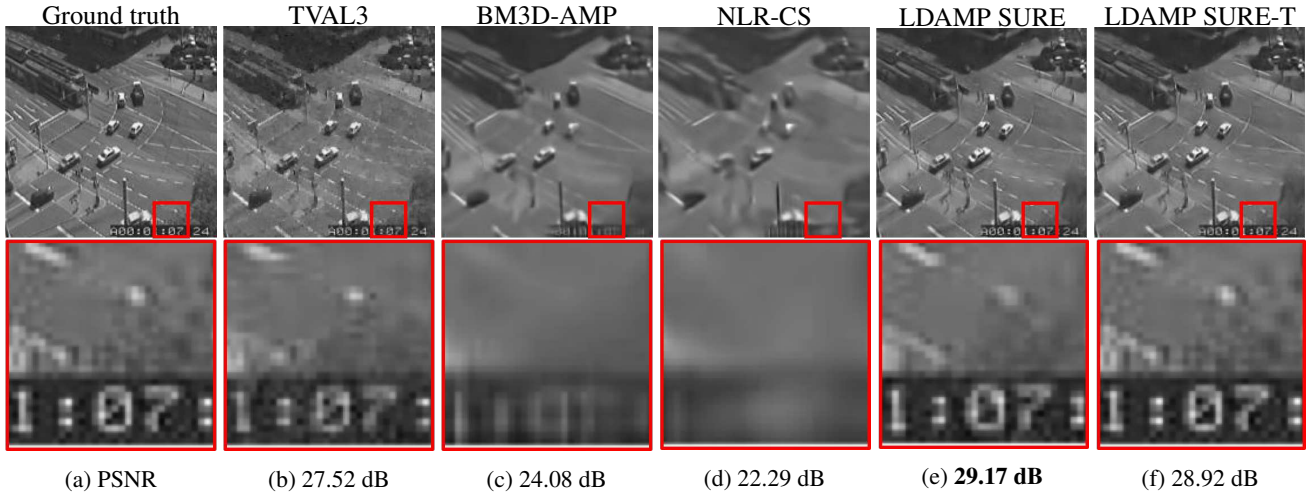
Figure 4: Reconstructions of 180×180 test image with CDP measurement matrix for $M/N = 0.15$ sampling rate.

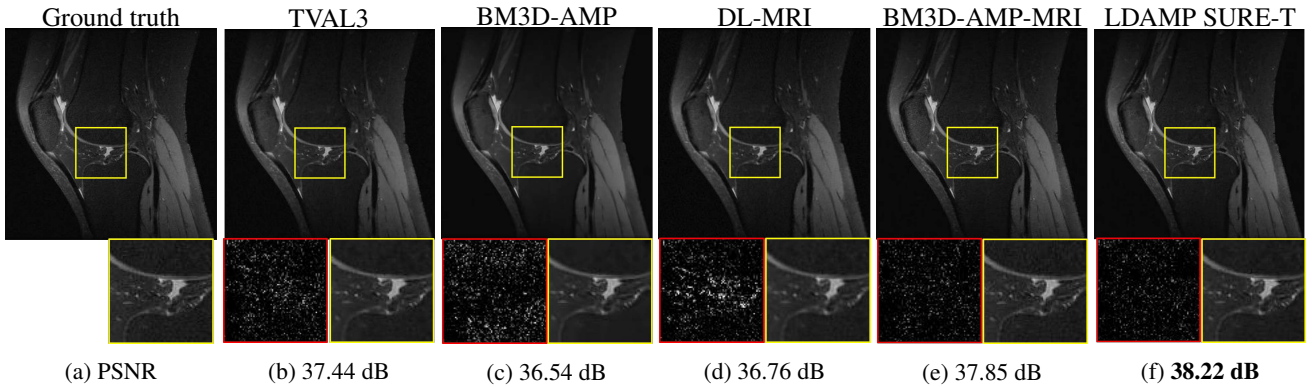| Ground truth	TVAL3	BM3D-AMP	DL-MRI	BM3D-AMP-MRI	LDAMP SURE-T
(a) PSNR	(b) 37.44 dB	(c) 36.54 dB	(d) 36.76 dB	(e) 37.85 dB	(f) **38.22 dB**

Figure 5: Reconstructions of 180×180 test image with CS-MRI measurement matrix for $M/N = 0.40$ sampling rate. Residual errors are shown in red boxes.

References

[1] Eirikur Agustsson and Radu Timofte. NTIRE 2017 challenge on single image super-resolution: Dataset and study. In *IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, 2017. 4

[2] Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle. Compressive Coded Aperture Spectral Imaging: An Introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, Nov. 2013. 1

[3] Amir Beck and Marc Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, Jan. 2009. 2

[4] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G Dimakis. Compressed sensing using generative models. In *International Conference on Machine Learning (ICML)*, pages 537–46, 2017. 2, 6

[5] A Bora, E Price, and A G Dimakis. AmbientGAN: Generative models from lossy measurements. In *International Conference on Learning Representations (ICLR)*, 2018. 2, 6

[6] Mark Borgerding, Philip Schniter, and Sundeep Rangan. AMP-inspired deep networks for sparse linear inverse problems. *IEEE Transactions on Signal Processing*, 65(16):4293–4308, 2017. 3

[7] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. A non-local algorithm for image denoising. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 60–65, 2005. 3

[8] Harold C Burger, Christian J Schuler, and Stefan Harmeling. Image denoising: Can plain neural networks compete with BM3D? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2392–2399, 2012. 2, 3

[9] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 39(2):277–299, 2015. 6

[10] E J Candes, J Romberg, and T Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, Jan. 2006. 1

[11] Kihwan Choi, Jing Wang, Lei Zhu, Tae-Suk Suh, Stephen Boyd, and Lei Xing. Compressed sensing based conebeam computed tomography reconstruction with a first-order method. *Medical Physics*, 37(9):5113–5125, Aug. 2010. 1

[12] 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. 2, 3

[13] Charles-Alban Deledalle, Samuel Vaiter, Jalal Fadili, and Gabriel Peyré. Stein Unbiased GrAdient estimator of the Risk (SUGAR) for multiple parameter selection. *SIAM Journal on Imaging Sciences*, 7(4):2448–2487, 2014. 6

[14] Weisheng Dong, Guangming Shi, Xin Li, Yi Ma, and Feng Huang. Compressive sensing via nonlocal low-rank regularization. *IEEE Transactions on Image Processing*, 23(8):3618–3632, 2014. 2, 6

[15] D L Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, Mar. 2006. 1

[16] D L Donoho, A Maleki, and A Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences (PNAS)*, 106(45):18914–18919, Nov. 2009. 2

[17] Ender M Eksioglu and A Korhan Tanc. Denoising AMP for MRI reconstruction: BM3D-AMP-MRI. *SIAM Journal on Imaging Sciences*, 11(3):2090–2109, 2018. 2, 5, 6

[18] Kevin Epperson, Anne Marie Sawyer, Michael Lustig, Marcus Alley, Martin Uecker, Patrick Virtue, Peng Lai, and Shreyas Vasanawala. Creation of Fully Sampled MR Data Repository for Compressed Sensing of the Knee. In *SMRT Conference*, 2013. 5

[19] Karol Gregor and Yann LeCun. Learning fast approximations of sparse coding. In *International Conference on International Conference on Machine Learning (ICML)*, pages 399–406, 2010. 2, 3

[20] Harshit Gupta, Kyong Hwan Jin, Ha Q Nguyen, Michael T McCann, and Michael Unser. CNN-Based Projected Gradient Descent for Consistent CT Image Reconstruction. *IEEE transactions on medical imaging*, pages 1–1, May 2018. 2

[21] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll. Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine*, 79(6):3055–3071, Nov. 2017. 2

[22] Viren Jain and Sebastian Seung. Natural image denoising with convolutional networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 769–776, 2009. 2, 3

[23] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, Sept. 2017. 2

[24] Diederik P. Kingma and Jimmy Lei Ba. ADAM: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015. 5

[25] Kuldeep Kulkarni, Suhas Lohit, Pavan K Turaga, Ronan Kerviche, and Amit Ashok. ReconNet: Non-iterative reconstruction of images from compressively sensed measurements. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 449–458, 2016. 2

[26] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. 2

[27] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In *International Conference on Machine Learning (ICML)*, pages 2965–74, 2018. 2, 6

[28] Chengbo Li, Wotao Yin, and Yin Zhang. Users guide for TVAL3: TV minimization by augmented Lagrangian and alternating direction algorithms. *CAAM report*, 20(46-47):4, 2009. 2, 6

[29] Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007. 1, 2

[30] 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 *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 416–423, July 2001. 4

[31] Christopher A Metzler, Arian Maleki, and Richard G Baraniuk. BM3D-AMP: A new image recovery algorithm based on BM3D denoising. In *IEEE International Conference on Image Processing (ICIP)*, pages 3116–3120, 2015. 3, 6

[32] Christopher A Metzler, Arian Maleki, and Richard G Baraniuk. From denoising to compressed sensing. *IEEE Transactions on Information Theory*, 62(9):5117–5144, 2016. 2, 3, 4, 6

[33] Christopher A Metzler, Arian Maleki, and Richard G Baraniuk. Learned D-AMP: Principled neural network based compressive image recovery. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1770–1781, 2017. 2, 3, 4

[34] Minh Phuong Nguyen and Se Young Chun. Bounded Self-Weights Estimation Method for Non-Local Means Image Denoising Using Minimax Estimators. *IEEE Transactions on Image Processing*, 26(4):1637–1649, Feb. 2017. 3

[35] Dongwon Park, Kwanyoung Kim, and Se Young Chun. Efficient module based single image super resolution for multiple problems. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 995–1003, 2018. 2, 3

[36] Lee C Potter, Emre Ertin, Jason T Parker, and Müjdat Cetin. Sparsity and Compressed Sensing in Radar Imaging. *Proceedings of the IEEE*, 98(6):1006–1020, May 2010. 1

[37] Sathish Ramani, Thierry Blu, and Michael Unser. Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms. *IEEE Transactions on Image Processing*, 17(9):1540–1554, 2008. 3

[38] Saiprasad Ravishankar and Yoram Bresler. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Transactions on Medical Imaging*, 30(5):1028–1041, 2011. 1, 2, 6

[39] Shakarim Soltanayev and Se Young Chun. Training deep learning based denoisers without ground truth data. In *Advances in Neural Information Processing Systems (NIPS)*, 2018. 2, 3, 6

[40] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre Antoine Manzagol. Stacked denoising autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research*, 11:3371–3408, Dec. 2010. 2, 3

[41] Y Wiaux, L Jacques, G Puy, A M M Scaife, and P Vandergheynst. Compressed sensing imaging techniques for radio interferometry. *Monthly Notices of the Royal Astronomical Society*, 395(3):1733–1742, May 2009. 1

[42] Junyuan Xie, Linli Xu, and Enhong Chen. Image denoising and inpainting with deep neural networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 341–349, 2012. 2, 3

[43] Kai Xu, Zhikang Zhang, and Fengbo Ren. LAPRAN: A scalable Laplacian pyramid reconstructive adversarial network for flexible compressive sensing reconstruction. In *European Conference on Computer Vision (ECCV)*, pages 491–507, 2018. 2

[44] Junfeng Yang and Yin Zhang. Alternating Direction Algorithms for ℓ_1-Problems in Compressive Sensing. *SIAM Journal on Scientific Computing*, 33(1):250–278, Jan. 2011. 2

[45] Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. Deep ADMM-Net for compressive sensing MRI. In *Advances in Neural Information Processing Systems (NIPS)*, pages 10–18, 2016. 2

[46] Xin Yuan, Patrick Llull, Xuejun Liao, Jianbo Yang, David J Brady, Guillermo Sapiro, and Lawrence Carin. Low-cost compressive sensing for color video and depth. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3318–25, 2014. 2

[47] Jian Zhang and Bernard Ghanem. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1828–1837, 2018. 2

[48] 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, 2017. 2, 3

[49] Lei Zhang, Wei Wei, Yanning Zhang, Fei Li, Chunhua Shen, and Qinfeng Shi. Hyperspectral compressive sensing using manifold-structured sparsity prior. In *IEEE International Conference on Computer Vision (ICCV)*, pages 3550–3558, 2015. 1, 2

[50] Lei Zhang, Wei Wei, Yanning Zhang, Chunna Tian, and Fei Li. Reweighted Laplace prior based hyperspectral compressive sensing for unknown sparsity. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2274–2281, 2015. 1, 2