SUREMAP: Predicting Uncertainty in CNN-based Image Reconstructions Using Stein's Unbiased Risk Estimate

#4467

Ruangrawee Kitichotkul rk22@stanford.edu
Christopher Metzler cmetzler@stanford.edu
Frank Ong franko@stanford.edu
Gordon Wetzstein gordon.wetzstein@stanford.edu
Department of Electrical Engineering, Stanford University

Motivation

- Convolutional neural networks (CNN) have emerged as a powerful tool for solving computational imaging reconstruction problems. However, it is challenging to know when they will work and, more importantly, when they will fail. This limitation is a major barrier to their use in safety-critical applications like medical imaging.
- Expected mean squared error (MSE), i.e. risk, is the gold standard for evaluating a compressive sensing (CS) reconstruction algorithm. However, computing the risk requires access to the ground truth image, which defeats the point of reconstruction in the first place.
- We use Stein's unbiased risk estimate (SURE) to estimate per-pixel MSE of images recovered from CS measurements with approximate message passing (AMP) framework in the form of heatmap.

Ground truth

Reconstructed image

Squared error

Related Work

- If the latent image lies in the range of a sufficiently expansive and invertible generative network, one can estimate the uncertainty from the network's latent variables [1].
- For probabilistic neural networks, one can sample from $p(\hat{\mathbf{x}}|\mathbf{y})$ and reason about the variance of the reconstruction [2].
- Edupuganti et al. used SURE to estimate the mean squared error associated with MRI reconstructions [3]. The method assumes that the difference between the true signal and an initial estimate follows a white Gaussian distribution, which does not hold in practice.

References

[1] Lynton Ardizzone, Jakob Kruse, Sebastian Wirkert, Daniel Rahner, Eric W Pellegrini, Ralf S Klessen, Lena Maier- Hein, Carsten Rother, and Ullrich Ko¨the, "Analyzing inverse problems with invertible neural networks," arXiv preprint arXiv:1808.04730, 2018.

[2] Jonas Adler and Ozan Oktem, "Deep posterior sampling: Uncertainty quantification for large scale inverse problems," in International Conference on Medical Imaging with Deep Learning–Extended Abstract Track, 2019.

[3] Vineet Edupuganti, Morteza Mardani, Shreyas Vasanawala, and John Pauly, "Uncertainty quantification in deep mri re- construction," IEEE Transactions on Medical Imaging, 2020.

Method

- Given a noisy signal $\mathbf{y} = \mathbf{x} + \eta$ where the noise follows a white Gaussian distribution i.e. $\eta \sim \mathcal{N}(0, \sigma^2\mathbf{I})$, an unbiased estimate of the MSE, $\frac{1}{n}\|\mathbf{x} - f(\mathbf{y})\|^2$ is SURE.

$$S(\mathbf{y}, f(\mathbf{y})) = \frac{1}{n}\|\mathbf{y} - f(\mathbf{y})\|^2 - \sigma^2 + \frac{2\sigma^2}{n}\text{div}_{\mathbf{y}}(f(\mathbf{y}))$$

- AMP is a simple iterative algorithm for reconstructing a signal from i.i.d. Gaussian measurements. The effective noise at every iteration follows a white Gaussian distribution.

$$\mathbf{r}_t = \mathbf{x} + \eta_t;\ \mathbf{x}, \mathbf{r}_t \in \mathbb{R}^n;\ \eta_t \sim \mathcal{N}(0, \sigma_\eta^2 \mathbf{I_n})$$

- When a closed form of the divergence is not available, it can be estimated with the following Monte-Carlo estimate.

$$\text{div}_{\mathbf{r}_T}(\hat{\mathbf{x}}) \approx \frac{1}{K}\sum_{k=1}^{K}\frac{1}{\epsilon}\mathbf{b}_k^T\left(f(\mathbf{r}_T + \epsilon\mathbf{b}_k) - f(\mathbf{r}_T)\right) \qquad \mathbf{b}_k \sim \mathcal{N}(0, \mathbf{I_n})$$

- Variable density AMP (VDAMP) is a recent extension to AMP which deals with variable density sampled Fourier measurements. The effective noise at each iteration follows a colored Gaussian distribution.

$$\mathbf{r}_t = \mathbf{x} + \eta_t;\ \mathbf{x}, \mathbf{r}_t \in \mathbb{R}^n;\ \eta_t \sim \mathcal{CN}(0, \mathbf{\Psi}^t \text{diag}(\tau_t)\,\mathbf{\Psi})$$

- An unbiased risk estimate for removing colored Gaussian noise $\eta_T \sim \mathcal{N}(0, \Sigma)$ is

$$S(\hat{\mathbf{x}}, \mathbf{r}_T) = \|\hat{\mathbf{x}} - \mathbf{r}_T\|^2 - \sum_{i=1}^{n}\tau_T^{(i)} \qquad \mathbf{u} = \mathbf{\Psi}\text{diag}\left(\tfrac{1}{2}\tau_t\right)^{-1}\mathbf{\Psi}^t\mathbf{r}_T$$

$$+ \frac{2}{n}\left(\text{div}_{\Re(\mathbf{u})}(\Re(\hat{\mathbf{x}})) + \text{div}_{\Im(\mathbf{u})}(\Im(\hat{\mathbf{x}}))\right)$$

- Since SURE equals MSE in expectation, we require many pixels to calculate an accurate SURE. To generate a SURE heatmap, we compute SURE of overlapping square patches and average the results.

Experimental Results

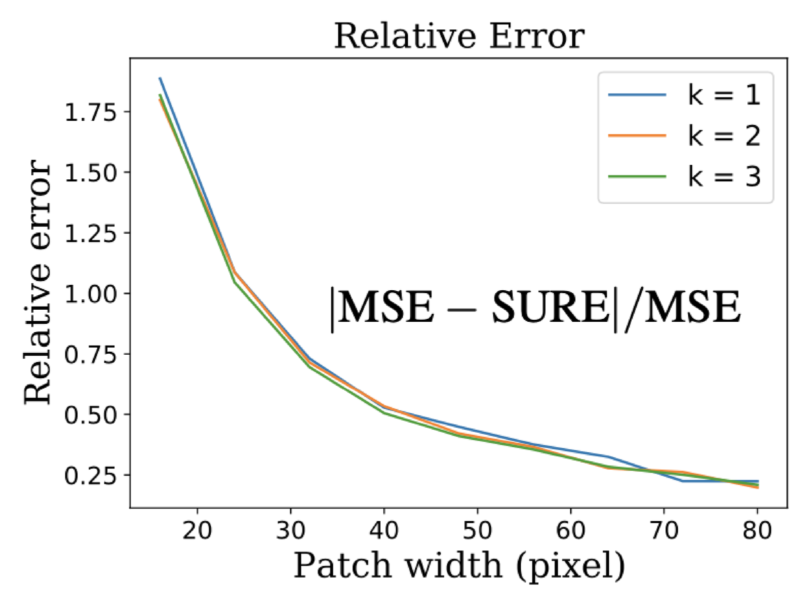

Fig. 1: Relative error of SURE heatmaps with respect to patch-average MSE heatmaps.

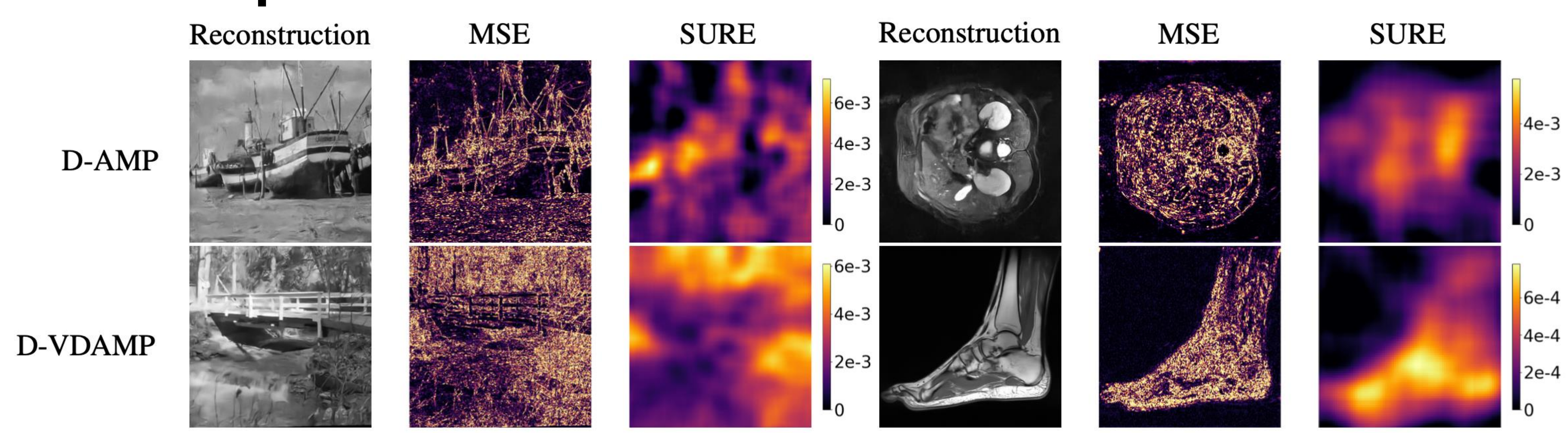

Fig. 2 SURE heatmaps and MSE heatmaps of CS reconstructions with D-AMP and D-VDAMP along with the reconstructed images.

- We generated SURE heatmaps of reconstructed images from the denoising-based versions of AMP and VDAMP, dubbed D-AMP and D-VDAMP respectively.
- Figure 1 compares the normalized absolute difference between the SURE estimate and the patch-average MSE of the image as one increases the patch sizes.
- We observe that the SURE heatmaps become more accurate as the patch size increase. Increasing the number of Monte-Carlo samples, K, has only a slight effect on the accuracy of the estimate.
- While smaller patch-sizes are higher resolution, larger patch sizes result in more accurate MSE estimates.
- We test our SURE heatmap generation method. For D-AMP, the sampling rate is 5%, and the SNRs are 23dB and 18dB for the natural image and the MR image. For D-VDAMP, the sampling rate is 25%, and the SNR is 20dB.
- Figure 2 shows the SURE heatmaps for D-AMP and D-VDAMP reconstructions using a patch size of 48x48 pixels. While somewhat low resolution, the shapes and magnitudes of the heatmaps closely follow the true pixelwise MSEs.