# Compressive SAR Image Recovery and Classification via CNNs

**Michael Wharton, Edward T. Reehorst, and Philip Schniter**

THE OHIO STATE UNIVERSITY

Supported by NSF grant IIP-0968910

**2019 Asilomar Conference on Signals, Systems, and Computers**

## Abstract

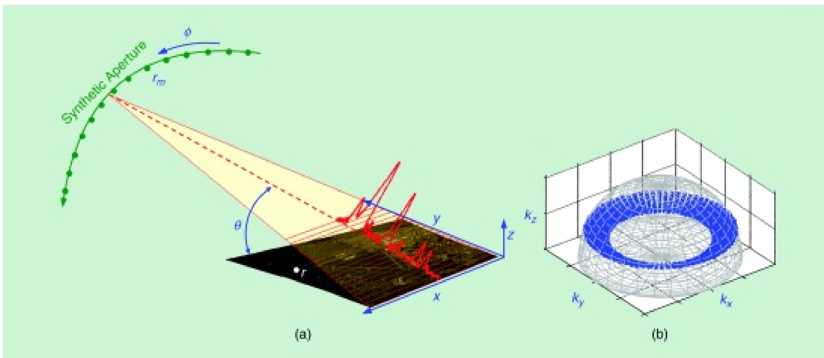
We consider synthetic aperture radar (SAR) image recovery and classification from sub-Nyquist samples, i.e., compressive SAR. Our approach is to first apply back-projection and then use a deep convolutional neural network (CNN) to de-alias the result. Importantly, our CNN is trained to be agnostic to the subsampling pattern. Relative to algorithmic SAR reconstruction approaches like LASSO, our CNN-based approach is much faster and more accurate, in terms of both MSE and classification error rate, on the MSTAR dataset.

## Linear Inverse Problems in Imaging

**Goal**: Recover $\boldsymbol{g}$ from noisy measurements $\boldsymbol{r} = \boldsymbol{A}\boldsymbol{g} + \boldsymbol{w}$, where $\begin{cases} \boldsymbol{g} \in \mathbb{C}^n : & \text{image} \\ \boldsymbol{A} \in \mathbb{C}^{m \times n} : & \text{known linear operator} \\ \boldsymbol{w} \in \mathbb{C}^m : & \text{noise} \end{cases}$

**Applications**:

1. deblurring
2. super-resolution
3. accelerated MRI
4. accelerated CT
5. microscopy (e.g., STORM)
6. synthetic aperture radar (SAR)



With active electronically steerable arrays (AESA), we can simultaneously image multiple scenes via sub-Nyquist sampling.

## SAR Measurement Model

With linear FM chirps, a uniform pulse repetition interval, and uniform sampling, we can approximate SAR measurements as noiseless, uniformly-spaced samples of the 2D Fourier transform on a *polar grid*:

$$\boldsymbol{r} = \boldsymbol{A}\boldsymbol{g} + \boldsymbol{w}.$$

## Traditional SAR

When these samples are taken at the Nyquist rate or higher, $\boldsymbol{A}$ has full column rank, and thus $\boldsymbol{g}$ can be accurately recovered using least-squares (LS):

$$\widehat{\boldsymbol{g}} = (\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathsf{H}}\boldsymbol{r}.$$

If $\boldsymbol{A}$ was orthonormal, the LS solution simplifies to back-projection:

$$\widehat{\boldsymbol{g}} = \boldsymbol{A}^{\mathsf{H}}\boldsymbol{r}.$$

This can be implemented by interpolating polar-format $\boldsymbol{r}$ onto a Cartesian grid and then applying a 2D-IFFT.

## Compressive SAR

### Compressive SAR

- We consider SAR image recovery and classification from sub-Nyquist samples [1].
- For this, we assume noiseless, subsampled 2D (Cartesian) Fourier measurements, i.e.,

$$\boldsymbol{r} = \boldsymbol{A}\boldsymbol{g} \quad \text{with} \quad \boldsymbol{A} = \boldsymbol{M}\boldsymbol{F}.$$

### Motivation

- With actively electronically steerable arrays (AESA), compressive SAR facilitates the simultaneous imaging of multiple scenes.
- Compressed returns are more efficient for storage and/or communication to the ground station.
- Certain anti-jamming approaches lead to sub-Nyquist sampling [1].

### Problem

- Since $\boldsymbol{A}$ is not full-column rank, it is impossible to accurately recover $\boldsymbol{g}$ without the use of additional prior information.
- Traditional estimates, such as those from back-projection or LS, contain aliasing artifacts.

### Baseline Approach

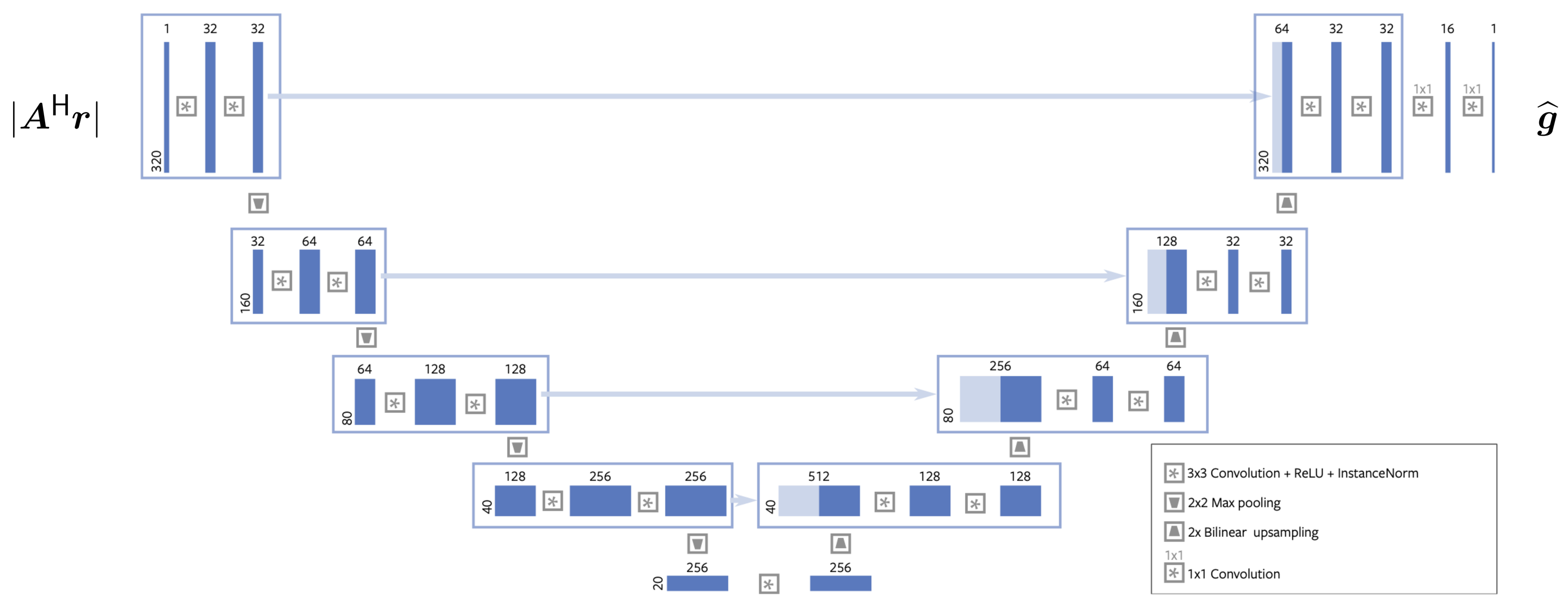
- Motivated by sparsity in the image domain, we consider LASSO (solved by FISTA [2]) as a baseline:

$$\widehat{\boldsymbol{g}} = \arg\min_{\boldsymbol{g}} \|\boldsymbol{g}\|_1 \text{ s.t. } \boldsymbol{A}\boldsymbol{g} = \boldsymbol{r}.$$

## Reconstruction U-Net

### De-aliasing network

- Our approach is to first use back-projection to form the aliased image $\boldsymbol{A}^{\mathsf{H}}\boldsymbol{r}$, and then to "de-alias" this image using a deep convolution neural network.
- We use a U-Net [3] because of its broad success in other image recovery problems.
- The input to the U-Net is the back-projection *magnitude*, and the output $\widehat{\boldsymbol{g}} \in \mathbb{R}^n_+$ is an estimate of $|\boldsymbol{g}|$.




### Training

- The U-Net $f_\theta(\cdot)$ is trained to minimize the $\ell_1$ loss

$$L(\theta) = \mathbb{E}_{\boldsymbol{g},\boldsymbol{M}}\left\{\left\|\boldsymbol{f}_\theta(|\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}\boldsymbol{g}|) - |\boldsymbol{g}|\right\|_1\right\},$$

where the expectation is taken over training images $\boldsymbol{g}$ and random sampling masks, $\boldsymbol{M}$, in $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{F}$.
- By training on many different masks, the learned network becomes agnostic to the sampling pattern.
- The use of $\ell_1$ loss (versus $\ell_2$ loss) is typical when training the U-Net.

## Image Reconstruction Results

### Experimental Setup

- We used the MSTAR dataset [4].
  - $17°$ inclination was used for training.
  - $15°$ inclination was used for testing.
- All ground-truth images were first center-cropped to size 128×128.
- We tested a variety of sampling rates $\delta \triangleq m/n$.
- We used a Linux server with 24 Intel Xeon(R) Gold 5118 CPUs and a Tesla V-100 GPU.

Reconstruction NMSE

| $\delta$ | FISTA | U-Net |
|---|---|---|
| 1/2 | −3.14 dB | **−9.59 dB** |
| 1/3 | −2.19 dB | **−8.36 dB** |
| 1/4 | −1.67 dB | **−7.75 dB** |
| 1/5 | −1.32 dB | **−7.25 dB** |
| 1/10 | −0.56 dB | **−6.24 dB** |

### Results

- The U-Net outperformed the baseline LASSO method for all tested sampling rates $\delta$ in both reconstruction NMSE (on the magnitude)

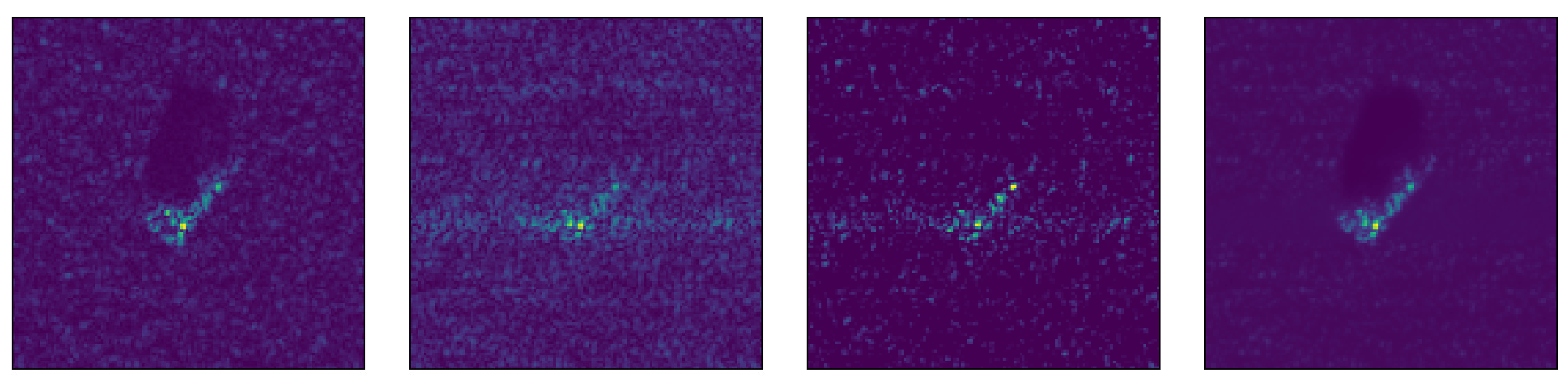
$$\text{NMSE}(\widehat{\boldsymbol{g}}, \boldsymbol{g}) = \frac{\left\||\widehat{\boldsymbol{g}}| - |\boldsymbol{g}|\right\|^2}{\|\boldsymbol{g}\|^2}$$

and computation time.

Computation Time

| FISTA | U-Net |
|---|---|
| 0.05917 sec | **0.00496 sec** |

- Example image reconstructions show that the U-Net tends to enhance the target's shadow and reduce image speckle:



Fully-sampled | Back-projection @ $\delta$=1/3 | FISTA (baseline) | U-Net (proposed)

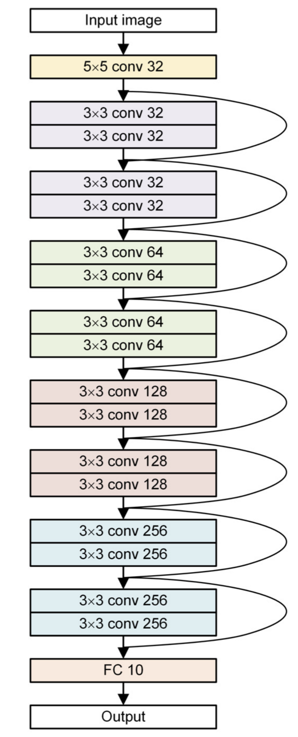
## Classifier for Automatic Target Recognition (ATR)

### Motivation

- SAR images are often used for Automatic Target Recognition (ATR) [5].
- In this case, classification accuracy is more important than image reconstruction NMSE.

### Classifier Network

- We used a ResNet-18 classification network [6] based on prior success with MSTAR data [7].
- The network was trained to minimize the standard cross-entropy loss.




## Compressive ATR Results

### Experimental Setup

1. First, a classifier was trained using noiseless, fully sampled images
   - It achieved $> 99\%$ accuracy.

   This classifier was then applied to classify the outputs of the LASSO and U-Net U-Net approaches to compressive SAR.
2. Next, a different classifier was trained using the reconstructed images output by LASSO and the U-Net at each sampling rate delta $\delta$.

Classifier trained on fully sampled data

| $\delta$ | FISTA | U-Net |
|---|---|---|
| 1/2 | 48.36 % | **75.96 %** |
| 1/3 | 39.71 % | **76.62 %** |
| 1/4 | 34.71 % | **73.87 %** |
| 1/5 | 28.91 % | **73.21 %** |
| 1/10 | 18.83 % | **63.16 %** |

Classifier trained on reconstructed images

| $\delta$ | FISTA | U-Net |
|---|---|---|
| 1/2 | 94.10 % | **99.38 %** |
| 1/3 | 89.42 % | **98.38 %** |
| 1/4 | 85.23 % | **97.80 %** |
| 1/5 | 80.02 % | **97.00 %** |
| 1/10 | 65.44 % | **91.10 %** |

### Results

- Classifiers trained on reconstructed images worked much better than the one trained on fully sampled images.
- U-Net reconstruction led to much better classification accuracy than FISTA reconstruction.
- With U-Net reconstruction at sampling rate $\delta = 1/2$, classification accuracy was essentially the same as on fully sampled data.

## Conclusion

### Contributions

- We proposed a novel method for compressive SAR image recovery that works by de-aliasing the back-projected images using a U-Net.
- Comparison to FISTA baseline:
  - The U-Net gave better performance in both NMSE and classification accuracy.
  - The U-Net ran $> 10\times$ faster.
- For compressive ATR, we observed that it was important to train the classifier on reconstructed images versus fully sampled images.

### Future Work

- We plan to jointly train both networks.
- We plan to test on more complicated datasets (e.g., ADTS [8]).

## References

1. V.M. Patel, G.R. Easley, D.M. Healy, and R. Chellappa, "Compressed synthetic aperture radar," *IEEE JSTSP*, vol. 4, no. 2, Apr 2010.
2. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sciences*, vol. 2, no. 1, Mar 2009.
3. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *MICCAI*, Nov 2015.
4. Ross, T. D., S. W. Worrell, V. J. Velten, J. C. Mossing, and M. L. Bryant, "Standard SAR ATR evaluation experiments using the MSTAR public release data set," in *Algorithms for Synthetic Aperture Radar Imagery V*, vol. 3370, pp. 566-573, 1998.
5. M. Wilmanski, C. Kreucher and J. Lauer, "Modern approaches in deep learning for SAR ATR," *SPIE*, vol. 9843, pp. 195-204, 2016.
6. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *IEEE CVPR*, Jun 2016.
7. H. Furukawa, "Deep learning for target classification from SAR imagery: Data augmentation and translation invariance, " *CoRR*, 2017. http://arxiv.org/abs/1708.07920
8. D. P. Morrison, A. C. Eckert, F. J. Shields, "Studies of advanced detection technology sensor (ADTS) data," *Algorithms for Synthetic Aperture Radar Imagery V*, vol. 2230, pp. 370-378, 1994.