COMPRESSIVE SAR IMAGE RECOVERY AND CLASSIFICATION VIA CNNS

Michael Wharton, Edward T. Reehorst, and Philip Schniter

Dept. ECE
The Ohio State University
Columbus, OH
{wharton.124, reehorst.3, schniter.1}@osu.edu

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 the basis pursuit (i.e., sparsity-based) approach to compressive SAR recovery, our CNN-based approach is faster and more accurate, in terms of both image recovery MSE and downstream classification accuray, on the MSTAR dataset.

1. INTRODUCTION

Synthetic aperture radar (SAR) uses a moving radar platform to transmit electromagnetic pulses and then uses the received echoes to estimate the scene reflectivity. We focus on spotlight SAR [1], where the radar continuously points at a given ground patch while transmitting and receiving pulses.

After demodulation, the sampled radar returns $\boldsymbol{r} \in \mathbb{C}^M$ can be expressed as [1]

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

where $\boldsymbol{g} \in \mathbb{C}^N$ is a vector of 2D scene reflectivity samples that we aim to recover, $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ is a linear operator, and $\boldsymbol{w}$ contains additive noise and clutter. With linear FM chirps, a uniform pulse repetition interval, and uniform sampling, $\boldsymbol{A}$ generates uniformly spaced samples along equi-spaced radial lines in 2D Fourier space, i.e., "polar format" samples [1].

When the samples $\boldsymbol{r}$ are taken at the Nyquist rate or above, $\boldsymbol{A}$ has full column rank, and thus $\boldsymbol{g}$ can be recovered using the least-squares approach

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

In fact, $\widehat{\boldsymbol{g}}$ from (1) would perfectly estimate $\boldsymbol{g}$ in the absence of noise. In practice, it is common to approximate (1) by interpolating the polar-format samples $\boldsymbol{r}$ onto a Cartesian grid and then applying a 2D IFFT to the result.

This work was supported in part by NSF Grant IIP-0968910

Recently, it has been proposed to sample $\boldsymbol{r}$ below the Nyquist rate (i.e., use $M < N$), which is known as *compressive SAR* [2, 3]. For example, one may choose to transmit and receive only a (possibly random) subset of the usual pulses, known as "slow-time subsampling." There are several motivations for doing this. For one, there is less information to store and/or transmit back to the ground station. Another is that the radar could simultaneously image multiple targets. Several other applications, include increased robustness to jamming, are discussed in [2].

In the compressive case, $\boldsymbol{A}$ is a fat matrix, in which case SAR image recovery is more challenging. In particular, $\boldsymbol{A}$ has a non-trivial nullspace, and all components of $\boldsymbol{g}$ in that nullspace are lost when collecting the measurements $\boldsymbol{r}$. Likewise, the inverse in (1) does not exist.

The traditional approach to compressive SAR recovery exploits image sparsity in an appropriate basis [2, 3], i.e., applies compressive sensing [4, 5]. For example, if $\boldsymbol{g}$ is sparse in the canonical basis (i.e., $\boldsymbol{g}$ is itself sparse), then one may attempt to recover $\boldsymbol{g}$ from $\boldsymbol{r}$ by solving the convex problem

$$\widehat{\boldsymbol{g}} = \arg\min_{\boldsymbol{g}} \|\boldsymbol{g}\|_1 \text{ s.t. } \|\boldsymbol{r} - \boldsymbol{A}\boldsymbol{g}\|^2 \leq M\sigma^2, \tag{2}$$

which is known as basis pursuit (BP) denoising [6]. This optimization problem (2) is convex and first-order algorithms like SPGL1 [7] can efficiently solve it. Still, these methods are computationally intensive for practical image sizes, and the sparsity model on which they are based may not fully exploit the structure of SAR images. Thus, one may wonder whether compressive SAR recovery can be performed using methods that are faster and/or more accurate.

2. CNN-BASED SAR IMAGE RECOVERY

We propose a convolutional neural network (CNN)-based approach to compressive SAR image recovery. In particular, we propose to first back-project the radar returns, yielding

$$\boldsymbol{z} \triangleq \boldsymbol{A}^{\mathsf{H}}\boldsymbol{r}. \tag{3}$$

With sub-Nyquist sampling, the back-projected scene $\boldsymbol{z}$ will be heavily aliased. We propose to then de-alias $\boldsymbol{z}$ using a

CNN. Among the plethora of CNN architectures, we chose a U-Net [8], because of its excellent performance in other image-recovery tasks [9].

For simplicity, we input only the magnitudes of the elements in $\boldsymbol{z}$ to the CNN. We do this because, in our experience, image phase information in $\boldsymbol{z}$ does not improve classification accuracy, at least with the MSTAR dataset that we used for our experiments. Altogether, our image-recovery approach can be summarized as

$$\widehat{\boldsymbol{g}} = \boldsymbol{f}(|\boldsymbol{z}|; \widehat{\boldsymbol{\theta}}), \tag{4}$$

where $\boldsymbol{f}(\cdot; \widehat{\boldsymbol{\theta}})$ is the CNN, $\widehat{\boldsymbol{\theta}}$ is a vector of trained CNN parameters, and $|\boldsymbol{z}|$ denotes the vector composed of the elementwise magnitudes of $\boldsymbol{z}$. The output $\widehat{\boldsymbol{g}}$ of our CNN is also non-negative, and thus should be considered as estimate of $|\boldsymbol{g}|$ rather than of complex-valued $\boldsymbol{g}$.

A similar approach was proposed for compressive magnetic resonance imaging (MRI) in [10], but—to the best of our knowledge—no CNN-based methods have been proposed for compressive SAR image recovery. However, CNNs have previously been proposed for other SAR tasks, such as image segmentation [11], image de-speckling [12], and automatic target recognition (ATR) [13, 14].

By training our CNN to de-alias the results of many *different* slow-time sub-sampling patterns (for a given sampling rate $\delta = M/N$), it learns to be agnostic to the specific choice of the sub-sampling pattern. This way, we do not need to retrain the CNN when the sub-sampling pattern changes. We did train a different CNN for each sampling rate $\delta \in \{1/2, 1/3, 1/4, 1/5, 1/10\}$, however. To learn the CNN parameters $\widehat{\boldsymbol{\theta}}$, we minimized ℓ_1 loss in the image space, i.e.,

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{t_1=1}^{T_1} \sum_{t_2=1}^{T_2} \left\| \boldsymbol{f}\big(|\boldsymbol{A}_{t_1}^{\mathsf{H}} \boldsymbol{A}_{t_1} \boldsymbol{g}_{t_2}|; \boldsymbol{\theta}\big) - |\boldsymbol{g}_{t_2}| \right\|_1, \tag{5}$$

using stochastic gradient descent. In (5), $\{\boldsymbol{g}_t\}$ are scene reflectivities from a training database and $\{\boldsymbol{A}_t\}$ are slow-time randomly sub-sampled Fourier matrices. We used the ℓ_1 loss, as opposed to the ℓ_2 loss, because it is a more typical choice when training CNNs to perform image recovery tasks [15].

3. CNN-BASED AUTOMATIC TARGET RECOGNITION

In the previous section, our goal was to recover the SAR image $\boldsymbol{g}$. Often, the recovered image is subsequently fed to an image classifier for automatic target recognition (ATR). In this case, the resulting classification performance is the primary metric of interest.

To evaluate our compressive SAR image recovery method from the perspective of ATR, we trained a ResNet-18 image classifier [16] to perform classification. Our choice of ResNet-18 was inspired by the excellent performance previously reported in [14]. For example, we found that a ResNet-18 gave 99.06 % classification accuracy with noiseless, fully sampled MSTAR images.

We experimented with two different approaches to training the ResNet classifier. We used either

1. noiseless, fully sampled MSTAR images, or
2. reconstructed MSTAR images produced by either BP or the U-Net, as described in Section 2, at a given value of $\delta = M/N$.

In both cases, we used the standard cross-entropy loss when training the classifier. As we will see in the next section, classification from compressive samples is much more accurate when the classifier is trained on compressively recovered images.

4. NUMERICAL RESULTS

For our numerical results, we assumed noiseless measurements, i.e.,

$$\boldsymbol{r} = \boldsymbol{A}\boldsymbol{g}. \tag{6}$$

with images $\boldsymbol{g}$ taken from the 10-class MSTAR dataset [17]. We used the 17°-inclination subset for training, which had $3\,671$ images and the 15°-inclination subset for testing, which had $3\,203$ images. Because the images are of various sizes, we first center-cropped them to size 128×128. To implement the polar-format Fourier operator $\boldsymbol{A}$, we quantized each point on each radial line to the nearest point on the 128×128 Cartesian grid. This allows us to approximate $\boldsymbol{A} \approx \boldsymbol{M}\boldsymbol{F}$, where $\boldsymbol{M}$ is a random masking operator and $\boldsymbol{F}$ is the 2D FFT operator. Various sampling ratios

$$\delta \triangleq \frac{M}{N} \tag{7}$$

were tested. Figure 1 shows an example of a mask at $\delta = 1/3$. Note the random subset of radial lines, with dense sampling across each line.

4.1. Image Recovery

As a baseline, we compare the proposed CNN-based method to BP recovery (i.e., equation (2) with $\sigma^2 = 0$) implemented using the SPGL1 algorithm [7]. We used the public MATLAB implementation of SPGL1[1] with default parameters.

Figure 1 shows an example of a Fourier-domain sub-sampling mask at sampling ratio $\delta = 1/3$, and Figure 2 shows an example of an MSTAR image. Figure 3 shows the result of backprojection, Figure 4 shows the SPGL1 recovery,

[1]MATLAB implementation of SPGL1 was downloaded from `https://www.cs.ubc.ca/~mpf/spgl1/index.html`

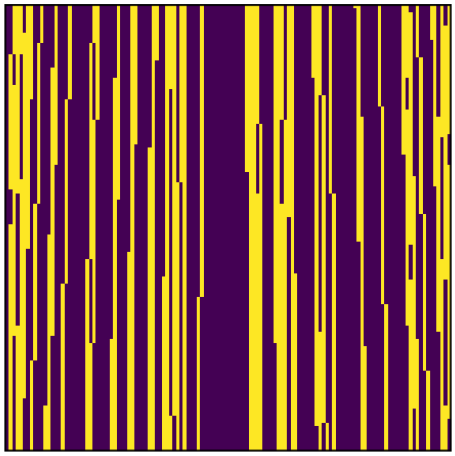

Fig. 1. Fourier-domain sampling mask at $\delta = 1/3$.

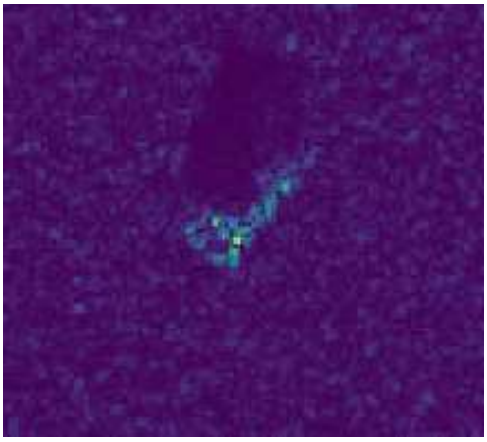

Fig. 2. Example of original MSTAR image $|\boldsymbol{g}|$.

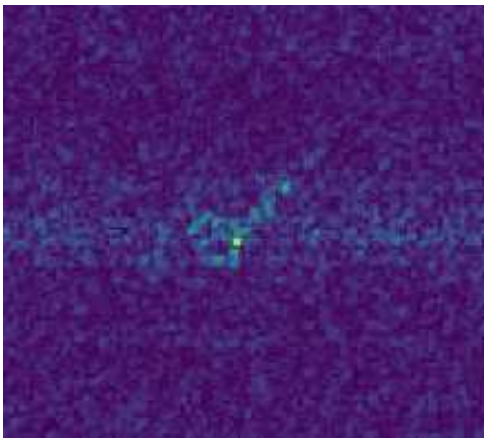

Fig. 3. Back-projection image $\boldsymbol{z}$ at $\delta = 1/3$.

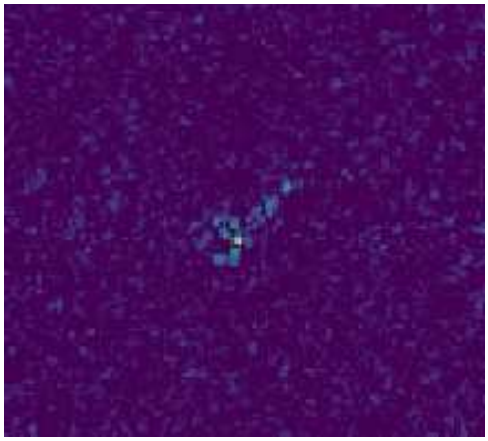

Fig. 4. SPGL1 reconstruction $\widehat{\boldsymbol{g}}$ at $\delta = 1/3$. Note that a lot of detail in the target has been lost relative to Fig. 2.

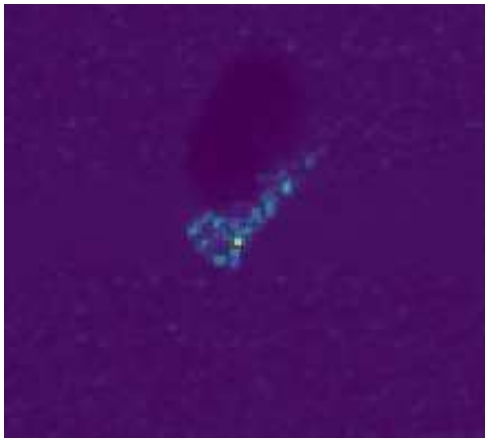

Fig. 5. U-Net reconstruction $\widehat{\boldsymbol{g}}$ at $\delta = 1/3$. Note that target details have been preserved relative to Fig. 2 while speckle artifacts have been suppressed.

Table 1. Average NMSE

δ	SPGL1	U-Net
1/2	-7.04 dB	$\mathbf{-10.63}$ **dB**
1/3	-4.68 dB	$\mathbf{-10.12}$ **dB**
1/4	-3.46 dB	$\mathbf{-8.43}$ **dB**
1/5	-2.69 dB	$\mathbf{-8.11}$ **dB**
1/10	-1.01 dB	$\mathbf{-6.92}$ **dB**

Table 2. Reconstruction time

δ	SPGL1	U-Net
1/2	2.64 sec	**0.00451 sec**
1/3	2.76 sec	**0.00496 sec**
1/4	2.94 sec	**0.00460 sec**
1/5	2.96 sec	**0.00445 sec**
1/10	3.25 sec	**0.00472 sec**

and Figure 5 shows the U-Net recovery. The SPGL1 recovery loses many details in the target that are visible in Figure 2, while the U-Net recovery preserves those details. It is interesting to observe that the U-Net recovery has suppressed most of the speckle artifacts that are present in the original MSTAR image Figure 2.

Table 1 shows the normalized mean squared error (NMSE) on the recovered image magnitudes, averaged over the test data $\{\boldsymbol{g}_t\}_{t=1}^T$, i.e.,

$$\text{NMSE} = \frac{1}{T}\sum_{t=1}^{T} \frac{\left\| \widehat{\boldsymbol{g}}_t - |\boldsymbol{g}_t| \right\|_2^2}{\|\boldsymbol{g}_t\|_2^2}, \tag{8}$$

where a different random subsampling mask was used for every test image. The table shows that the proposed U-Net recovery method greatly outperformed SPGL1 for all tested sub-sampling rates δ. We attribute the relatively poor performance of SPGL1 to the large amounts of speckle present in the original MSTAR images (see, e.g., Figure 2), which detract from the sparsity of the image.

Table 2 shows the reconstruction time on a Linux server with 24 Intel Xeon(R) Gold 5118 CPUs and a single Tesla V-100 GPU. The table shows that the proposed method ran more than 500 times faster than SPGL1.

4.2. Automatic Target Recognition

Table 3 shows test classification accuracy at different sub-sampling ratios δ for the ResNet classifier that was trained on noiseless, fully sampled MSTAR images. This table shows that the U-Net-reconstructed images lead to much better classification accuracy than the SPGL1-reconstructed images, but in both cases the classification accuracy is far from the 99% achieved by the ResNet on fully sampled test images (i.e., non-compressive SAR).

Table 3. Average test classification accuracy using the ResNet classifier trained on fully sampled images

δ	SPGL1	Proposed
1/2	86.11 %	**87.42 %**
1/3	76.58 %	**88.29 %**
1/4	65.41 %	**86.73 %**
1/5	54.70 %	**85.76 %**
1/10	32.19 %	**74.37 %**

Table 4. Average test classification accuracy using the ResNet classifier trained on reconstructed images

δ	SPGL1	U-Net
1/2	96.85 %	**99.38 %**
1/3	94.01 %	**98.38 %**
1/4	90.67 %	**97.80 %**
1/5	86.58 %	**97.00 %**
1/10	70.12 %	**91.10 %**

Table 4 shows classification accuracy at different sub-sampling ratios δ for the ResNet classifiers trained on reconstructed images. Note that a different classifier was trained for SPGL1 and for the U-Net at each sub-sampling rate δ. This table shows that, as before, U-Net image reconstruction leads to much better classification accuracy than SPGL1 image reconstruction. However, differently from before, Table 4 shows that the U-Net recoveries from compressive SAR lead to classification accuracies on par with fully sampled SAR. In fact, at a sampling rate of $\delta = 1/2$, U-Net recovery yields a classification accuracy of 99.38%, which is slightly better than that achieved in the fully sampled case.

5. CONCLUSION

In this paper, we proposed a novel approach to compressive SAR image recovery that used a convolutional neural network (CNN) to de-alias the back-projection image estimate. Numerical experiments with the MSTAR dataset showed that our approach significantly outperformed BP—the standard compressed sensing approach—in both runtime and mean-squared error. To evaluate the quality of image recovery for subsequent use in automatic target recognition, we trained a second CNN to classify the reconstructed images. There we found that the accuracy of classifying compressively sampled images with with CNN-based image reconstruction and sampling rate $\delta = 1/2$ was on par with the accuracy of classifying fully sampled images.

6. REFERENCES

[1] D. C. Munson, J. D. O'Brien, and W. K. Jenkins, "A tomographic formulation of spotlight-mode synthetic aperture radar," in *Proc. IEEE*, vol. 71, pp. 917–925, Aug. 1983.

[2] V. M. Patel, G. R. Easley, J. Dennis M. Healy, and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Sel. Topics Signal Process.*, vol. 4, pp. 244–254, Apr. 2010.

[3] L. C. Potter, E. Ertin, J. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, pp. 1006–1020, June 2010.

[4] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21–30, Mar. 2008.

[5] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, "Introduction to compressed sensing," in *Compressed Sensing: Theory and Applications* (Y. C. Eldar and G. Kutyniok, eds.), Cambridge Univ. Press, 2012.

[6] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientific Comput.*, vol. 20, pp. 33–61, 1998.

[7] E. van den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM J. Scientific Comput.*, vol. 31, no. 2, pp. 890–912, 2008.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Intl. Conf. Med. Image Comput. & Computer-Assisted Intervention*, pp. 234–241, 2015.

[9] J. Zbontar, F. Knoll, A. Sriram, M. J. Muckley, M. Bruno, A. Defazio, M. Parente, K. J. Geras, J. Katsnelson, H. Chandarana, Z. Zhang, M. Drozdzal, A. Romero, M. Rabbat, P. Vincent, J. Pinkerton, D. Wang, N. Yakubova, E. Owens, C. L. Zitnick, M. P. Recht, D. K. Sodickson, and Y. W. Lui, "fastMRI: An open dataset and benchmarks for accelerated MRI," *arXiv:1811.08839*, 2018.

[10] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Sleu, "Deep learning for undersampled MRI reconstruction," *Physics in Medicine and Biology*, vol. 63, June 2018.

[11] Z. Zhang, H. Wang, F. Xu, and Y. Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification," *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 55, pp. 7177–7188, Dec. 2017.

[12] P. Wang, H. Zhang, and V. M. Patel, "SAR image despeckling using a convolutional neural network," *IEEE Signal Process. Lett.*, vol. 24, pp. 1763–1767, May 2017.

[13] D. A. E. Morgan, "Deep convolutional neural networks for ATR from SAR imagery," in *Proc. SPIE Defense and Security*, May 2015.

[14] H. Furukawa, "Deep learning for target classification from SAR imagery: Data augmentation and translation invariance," *arXiv:1708.07920*, Aug. 2017.

[15] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comp. Imag.*, vol. 3, no. 1, pp. 47–57, 2016.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv:1512.03385*, 2015.

[17] T. D. Ross, 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 *Proc. SPIE*, vol. 3370, Sept. 1998.