Compressive SAR Image Recovery and Classification via CNNs

Michael Wharton, Edward T. Reehorst, and Philip Schniter

Support by NSF grant IIF-0968910

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 algebraic 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

Motivation

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 to de-alias the result. Importantly, our CNN is trained to be agnostic to the subsampling pattern. Relative to algebraic 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.

Compressive SAR

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

\[ \mathbf{g} = \mathbf{A} \mathbf{f} + \mathbf{n} \]

where \( \mathbf{g} \) is the sampled data, \( \mathbf{A} \) is the known linear operator, \( \mathbf{f} \) is the unknown object, and \( \mathbf{n} \) is additive white Gaussian noise.

Traditional SAR

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

\[ \hat{\mathbf{f}} = \mathbf{A}^+ \mathbf{g} \]

where \( \mathbf{A}^+ \) is the Moore-Penrose pseudoinverse of \( \mathbf{A} \).

If \( \mathbf{A} \) is orthonormal, the LS solution simplifies to back-projection:

\[ \hat{\mathbf{g}} = \mathbf{A}^+ \mathbf{g} \]

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

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.,

\[ \mathbf{r} = \mathbf{A} \mathbf{g} + \mathbf{n} \]

where \( \mathbf{r} \) is the observed data, \( \mathbf{n} \) is additive white Gaussian noise, and \( \mathbf{A} \) is a known linear operator.

Motivation

With actively electronically steerable arrays (AESAs), 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 [2].

Problem

Because \( \mathbf{A} \) is not full-column rank, it is impossible to accurately recover \( \mathbf{f} \) without the use of additional 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:

\[ \hat{\mathbf{f}} = \text{arg min}_{\mathbf{f}} \frac{1}{2} \| \mathbf{A} \mathbf{f} - \mathbf{r} \|^2 + \lambda \| \mathbf{f} \|_1 \]

where the expectation is taken over training images \( \mathbf{r} \). If \( \mathbf{A} \) is not full-column rank, it is impossible to accurately recover \( \mathbf{f} \) without the use of additional information.

Traditional estimates, such as those from back-projection or LS, contain aliasing artifacts.

Reconstruction U-Net

De-aliasing network

Our approach is to first use back-projection to form the aliased image \( \mathbf{A}^+ \mathbf{r} \), 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 \( \hat{\mathbf{g}} \in \mathbb{R}^n \) is an estimate of \( \mathbf{g} \).

Image Reconstruction Results

Experimental Setup

We used the MSTAR dataset [4].

1. Training images were used for training.

2. Test images were used for testing.

3. All ground-truth images were first center-cropped to size 128x128.

4. We tested a variety of sampling rates \( \delta = 1/5, 1/4, 1/3, 1/2 \).

5. We used a Linux server with 24 Intel Xeon(R) Gold 5118 CPUs and a Tesla V100 GPU.

Results

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

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>Reconstructed Magnitude</th>
<th>U-Net</th>
<th>LASSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/5</td>
<td>0.125</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>1/4</td>
<td>0.175</td>
<td>0.13</td>
<td>0.18</td>
</tr>
<tr>
<td>1/3</td>
<td>0.250</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td>1/2</td>
<td>0.312</td>
<td>0.22</td>
<td>0.25</td>
</tr>
</tbody>
</table>

and computation time.

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

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.

Classifier trained on fully sampled data

<table>
<thead>
<tr>
<th>Classifier</th>
<th>NMSE</th>
<th>ATR Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>FISTA</td>
<td>0.20</td>
<td>99%</td>
</tr>
<tr>
<td>U-Net</td>
<td>0.10</td>
<td>99%</td>
</tr>
</tbody>
</table>

Comparison to FISTA baseline:

The U-Net gave better performance in both NMSE and classification accuracy.

The U-Net can run >10x 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


