A Novel Approach for SAR Target Detection Based on Unsupervised Complex-Valued Extreme Learning Machine

Qinglong Hua¹, Yun Zhang¹, and Yicheng Jiang¹

¹Harbin Institute of Technology

August 1, 2023

Abstract

Strong clutter seriously affects target-of-interest detection in synthetic aperture radar (SAR) images. This letter proposes an unsupervised target detection method (U-TDM) based on a complex-valued extreme learning machine (CV-ELM), the essence of which is to transform the problem of target detection into a pixel binary classification problem. The SAR image is first divided into several unlabeled patches, and fuzzy c-means (FCM) is used to construct the reference target patch set and the clutter patch set. Based on these two patch sets, CV-ELM is used to classify the neighboring patch of the pixel to be detected. Since the pixel intensity and distribution of target-of-interest and clutter are different, unsupervised pixel classification could be realized without ground-truth through U-TDM. Experimental results on GF-3 data and Sentinel-1 data show the efficiency of the proposed method in target detection with a heterogeneous clutter environment.

Hosted file


Hosted file

A Novel Approach for SAR Target Detection
Based on Unsupervised Complex-Valued Extreme Learning Machine

Qinglong Hua,1 Yun Zhang,1 YiCheng Jiang1
1School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, 150001, China
Email: zhangyunhit@hit.edu.cn

Strong clutter seriously affects target-of-interest detection in synthetic aperture radar (SAR) images. This letter proposes an unsupervised target detection method (U-TDM) based on a complex-valued extreme learning machine (CV-ELM), the essence of which is to transform the problem of target detection into a pixel binary classification problem. The SAR image is first divided into several unlabeled patches, and fuzzy c-means (FCM) is used to construct the reference target patch set and the clutter patch set. Based on these two patch sets, CV-ELM is used to classify the neighboring patch of the pixel to be detected. Since the pixel intensity and distribution of target-of-interest and clutter are different, unsupervised pixel classification could be realized without ground-truth through U-TDM. Experimental results on GF-3 data and Sentinel-1 data show the efficiency of the proposed method in target detection with a heterogeneous clutter environment.

Introduction: Synthetic aperture radar (SAR) is an active microwave sensor that can perform all-time and all-weather imaging without being affected by light and climate [1]. SAR images are usually composed of target-of-interest and clutter. The strong sea clutter or ground clutter increases the difficulty of target detection. The quality of target detection dramatically affects the performance of subsequent SAR image processing. Accurate detection results are conducive to more accurately identifying the target’s category and superstructure information.

After decades of development, many detection methods for SAR images have been proposed. The classic constant false alarm rate (CFAR) detector [2], [3] relies on contrast or brightness features to search for targets, which may result in a large number of clutter false alarms in the complex background of strong clutter. Moreover, since moving targets and stationary clutter have different Doppler parameters, spectral filter and time-frequency analysis could be used to detect moving targets in a single-channel SAR system. Huang et al. [4] proposed a time-Doppler chirp-filtering (TDCC) filter for moving target detection and parameter estimation with the airborne system. The multi-channel SAR system could increase the spatial information of the echo and improve the ability to observe moving targets [5]. These methods have higher requirements for the consistency of system channels. The extended fractal (EF) method [6] is another representative research strategy in the next.

In U-TDM, fuzzy c-means (FCM) [13] is first used to construct the reference target and clutter patch set. Then target patches and clutter patches randomly selected from these two patch sets are classified by the CV-ELM together with the neighborhood patch of the pixel to be detected. Since the target and clutter are different in pixel intensity and distribution, unsupervised pixel classification could be achieved without actual labels or ground-truth. Experiments on GF-3 and Sentinel-1 data verify the effectiveness of the U-TDM.

U-TDM: The implementation details of the proposed unsupervised target detection method are shown in Fig. 1. Next, we elaborate on the U-TDM from three perspectives: reference patch sets, unsupervised strategy, and CV-ELM.

Construction of Reference Patch Sets: Before target detection in a SAR image S, two reference patch set, target patch set SN and clutter patch set NC, are necessary to be constructed. First, S is uniformly cropped into a series of patches of size (2d + 1) × (2d + 1), where d is the scale factor. These unlabeled patches are classified into target and clutter by FCM. The target patches with membership greater than the threshold Tc,of constitute the target patch set SN = {T1, T2, ..., TN}, and the clutter patches with membership greater than the threshold Tc,of form the clutter patch set NC = {C1, C2, ..., CN}. SN and NC are the number of the target patch set and the clutter patch set, respectively. It should be noted that only background clutter is included in the clutter patch. The target patch contains both the target-of-interest and the background clutter. The reference patch sets could be regarded as the assisted knowledge used in U-TDM. Based on the reference patch sets, construct an unsupervised strategy in the next.

Unsupervised Strategy: Consider a complex SAR image S ∈ C2M×N. M is the number of rows, and N is the number of columns. s(m, n) represents the m-th row and n-th column of S. Define Q ∈ C2M×N as the ground-truth of S. Define q(m, n) as the m-th row and n-th column of Q. q(m, n) also represents the category of s(m, n). If the category is the target, then q(m, n) = 1 + i1, otherwise q(m, n) = 0.

For the pixel s(m, n) in the SAR image S, it is difficult to effectively determine its category based on the information of a single pixel. Generally, the neighborhood of the pixel s(m, n) has a high degree of correlation with s(m, n), so it is feasible to use the neighborhood information of s(m, n) to determine the category of s(m, n). Define the neighborhood of s(m, n) as NH(m, n).

\[ \text{NH}(m, n) = \{s(m - d, n - d), \ldots, s(x, y), \ldots, s(m + d, n + d)\} \]

where \( s(x, y) \) is the \( j \)-th pixel of \( \text{NH}(m, n) \), \( 1 < m < M \), \( 1 < n < N \). The number of pixels in \( \text{NH}(m, n) \) is \( N = (2d + 1)^2 \).

Define two complex-valued functions \( f^k(\cdot) \), \( k = 1, 2 \). When vectorized \( \text{NH}(m, n) \) is input, the output of \( f^1(\cdot) \) is \( q(m, n) \). When the vectorized reference target patch is input, the output of \( f^2(\cdot) \) is 1. When the vectorized reference clutter patch is input, the output of \( f^2(\cdot) \) is 0. Then the function \( f^k(\cdot) \) could be expressed as

\[ f^k_{m,n} = \min_{f^k(\cdot)} \|L_{k}^{f^k_{m,n}}(\text{NH}(m, n)) - f^k(\text{vec}(P^k_{m,n}(m, n)))\|^2_F \]

where \( [P^k_{m,n}]_{C^2(2d+1)^2}^{2M \times N} \) is the input of \( f^k_{m,n}(\cdot) \), and \( \{L^k_{m,n}\}_{C^2(2d+1)^2}^{2M \times N} = \{q(m, n), 1 + i1, \ldots, 1 + i1\} \) is the output of \( f^k_{m,n}(\cdot) \). \( \{P^k_{m,n}\}_{C^2(2d+1)^2}^{2M \times N} = \{\text{NH}(m, n), C_1, \ldots, C_N\} \) is the input of \( f^2_{m,n}(\cdot) \). \( \|\cdot\|^2_F \) denotes the Frobenius norm. \( N \) represents the number of patches randomly selected from the reference patch sets. \( \text{vec}(P^k_{m,n}) \) represents vectorized \( P^k_{m,n} \). Each row of data in \( P^k_{m,n} \) is connected end to end to obtain a single row vector.

Based on the reference patches, the classification problem of SAR image pixels is converted to solve a group of functions as formula (6). The classification result of the SAR image \( S \) could be obtained by integrating \( \hat{Q}^1 \) and \( \hat{Q}^2 \), that is, \( \hat{Q} = \hat{Q}^1 \odot \hat{Q}^2 \) represents the Hadamard product.
In supervised learning algorithms, knowing the value of \( q(m,n) \) in advance is necessary before learning and solving the formula (6). But in the application, it’s not easy to obtain the true value of \( q(m,n) \) directly. Therefore, this letter adopts an unsupervised training strategy. In this strategy, the actual value of \( NH(m,n) \) belongs to the target or clutter, \( q(m,n) \) in \( L^1 \) is set to 0. \( q(m,n) \) in \( L^1 \) is set to 1 + 1j. Next, the rationale of this strategy is explained by taking \( f_{m,n}^j(\cdot) \), \( P^1 \), and \( L^2 \) as examples. If the central point pixel \( s(m,n) \) of \( NH(m,n) \) belongs to the target of interest, the actual value of \( q(m,n) \) in \( L^1 \) is 1 + 1j. Since the target and clutter have different intensities in pixel intensity and distribution, \( f_{m,n}^j \) can classify \( NH(m,n) \) and reference clutter patches randomly selected from \( NC \) well. If the central point pixel \( s(m,n) \) of \( NH(m,n) \) belongs to clutter, the actual value of \( q(m,n) \) in \( L^1 \) is 0. Although \( q(m,n) \) is forcibly set to 1 + 1j, \( f_{m,n}^j \) is limited in fitting ability and there are enough clutter patches. \( f_{m,n}^j \) has difficulty separating \( NH(m,n) \) from reference clutter patches and would tend to classify \( NH(m,n) \) into the clutter category. Therefore, even if the label \( q(m,n) \) is set to 1 + 1j, there is a chance that \( s(m,n) \) belongs to the clutter, the correct classification result could be wrong. Similarly, setting \( q(m,n) \) to 0 in \( L^1 \) is also reasonable.

CV-ELM: For the above mapping function \( f_{m,n}^j(\cdot) \), CV-ELM proposed in this letter is used to solve it, which is a complex-valued version of ELM. ELM is one of the outstanding models in the field of pattern recognition. Its network structure is similar to traditional neural networks, and its learning efficiency is better. ELM provides a faster training convergence and does not require repeated iterations to adjust weights. It has the advantages of fast learning speed and good generalization performance. However, ELM could only process real-valued data. For complex-valued input \( x \), the activation process of \( f(\cdot) \) is given by

\[
\sigma(z) = \text{tanh}(i \arg(z))
\]

The loss function of CV-ELM is the quadratic norm of the difference between the expectation and output, expressed as follows.

\[
\mathbf{H}\mathbf{f} = \mathbf{L}
\]

where \( \mathbf{L} \) is the expectation or label set by unsupervised strategy. For \( \mathbf{L} \), \( \mathbf{L} = [0, 1, 1, ..., 1] \). For \( \mathbf{L} \), \( \mathbf{L} = [1, 1, 1, 0, ..., 0] \). \( \mathbf{H} \in \mathbb{C}^{N_L \times N_L} \) represents the output matrix of the hidden layer. Let \( N_L = N + 1 \), then \( \mathbf{H} \) could be expanded into the following form.

\[
\mathbf{H} = \begin{bmatrix}
\sigma(X_1W_1 + b_1) & \cdots & \sigma(X_1W_L + b_L) \\
\vdots & \ddots & \vdots \\
\sigma(X_{N_L}W_1 + b_1) & \cdots & \sigma(X_{N_L}W_L + b_L)
\end{bmatrix}
\]

Then the output weight \( \mathbf{f} \) is

\[
\mathbf{f} = \mathbf{H}^\dagger \mathbf{L}
\]

where \( \mathbf{H}^\dagger \) is the Moore-Penrose generalized inverse.

**Experimental Results:** The proposed method is compared with two representative methods, CFAR and EF. The GF-3 SAR image containing sea clutter and a SAR dataset containing 207 ship targets are used to test the effectiveness of the proposed method in different scenes. Three indicators of missed-detection probability \( P_m \) [15], recall \( P_r \), and precision \( P_p \) [16] are used to measure the performance of the proposed method. \( P_m \) represents the proportion of target pixels that are incorrectly classified as clutter pixels to all the target pixels. \( P_r \) is the proportion of target pixels correctly classified to all target pixels. \( P_p \) represents the proportion of target pixels that are correctly classified to all the pixels classified as target pixels. For the experimental parameters, the node number of the hidden layer in CV-ELM is 51, the size of \( NH(m,n) \) is 3x3, and \( N_{l} = 400 \). The DoF is 0.8. Conventionally, the degree of freedom (DoF) of CV-ELM is twice that of a counterpart ELM. In order to make a fair comparison, an ELM is designed to have the same DoF as the CV-ELM. The node number of the hidden layer in ELM is 102.

**Performance Comparison on A GF-3 SAR Image:** The processing results of different methods for a GF-3 image are shown in Fig. 2. The GF-3 image has one ship target and strong and heterogeneous sea clutter. The \( P_m \), \( P_r \), and \( P_p \) performance comparisons are shown in Table 1. According to the qualitative analysis in Fig. 2, the traditional method of CFAR could detect most of the pixels in ship targets. Still, it also incorrectly detects some strong clutter as targets. By comparing the original image in Fig. 2(a) and the detection result of the EF in Fig. 2(d), strong
Conclusion: In this letter, we propose a U-TDM based on an unsupervised strategy and CV-ELM for SAR target detection. CV-ELM converts the target detection problem into a pixel binary classification problem, which could fully exploit the characteristic differences in amplitude and phase between the target and clutter. In U-TDM, the weights of CV-ELM are solved by an unsupervised strategy based on reference patch sets, which avoids the acquisition difficulty of ground-truth. Experiments show that the proposed method could improve target detection performance compared with traditional methods. In the future study, the proposed method will be used to detect targets in a radio-frequency interference background.

References


Table 1. Performance Comparison on GF-3 SAR Image (%)

<table>
<thead>
<tr>
<th></th>
<th>CFAR</th>
<th>EF</th>
<th>U-TDM (ELM)</th>
<th>U-TDM (CV-ELM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{m}$</td>
<td>49.72</td>
<td>47.73</td>
<td>7.10</td>
<td>5.68</td>
</tr>
<tr>
<td>$P_{f}$</td>
<td>50.28</td>
<td>52.27</td>
<td>92.90</td>
<td>94.32</td>
</tr>
<tr>
<td>$P_{p}$</td>
<td>25.00</td>
<td>41.16</td>
<td>68.27</td>
<td>69.60</td>
</tr>
</tbody>
</table>

Table 2. Performance Comparison on the SAR Ship Dataset

<table>
<thead>
<tr>
<th></th>
<th>Correctly detected target</th>
<th>False alarm</th>
<th>Missing alarm</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFAR</td>
<td>153</td>
<td>64</td>
<td>54</td>
</tr>
<tr>
<td>EF</td>
<td>120</td>
<td>84</td>
<td>87</td>
</tr>
<tr>
<td>U-TDM (ELM)</td>
<td>195</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>U-TDM (CV-ELM)</td>
<td>198</td>
<td>7</td>
<td>9</td>
</tr>
</tbody>
</table>

Fig 2 Results of target detection for GF-3 image. (a) Original image. (b) Ground-truth. (c) CFAR processed image. (d) EF processed image. (e) U-TDM with ELM processed image. (f) U-TDM with CV-ELM processed image.

Fig 3 Result of ship detection for Sentinel-1 SAR images. The correctly detected targets, the false alarms, and the missed alarms are indicated by the green, red, and blue rectangle boxes, respectively.