



Article

Siamese Network Ensembles for Hyperspectral Target Detection with Pseudo Data Generation

Xiaodian Zhang , Kun Gao * , Junwei Wang , Zibo Hu , Hong Wang and Pengyu Wang

Key Laboratory of Photoelectronic Imaging Technology and System, Ministry of Education of China, Beijing Institute of Technology, Beijing 100081, China; xdz@bit.edu.cn (X.Z.); wjw@bit.edu.cn (J.W.); zibohu@bit.edu.cn (Z.H.); 3120185321@bit.edu.cn (H.W.); 3120210542@bit.edu.cn (P.W.)

* Correspondence: gaokun@bit.edu.cn

Abstract: Target detection in hyperspectral images (HSIs) aims to distinguish target pixels from the background using knowledge gleaned from prior spectra. Most traditional methods are based on certain assumptions and utilize handcrafted classifiers. These simple models and assumptions' failure restrict the detection performance under complicated background interference. Recently, based on the convolutional networks, many supervised deep learning detectors have outperformed the traditional methods. However, these methods suffer from unstable detection, heavy computation burden, and optimization difficulty. This paper proposes a Siamese fully connected based target detector (SFCTD) that comprises nonlinear feature extraction modules (NFEMs) and cosine distance classifiers. Two NFEMs, which extract discriminative spectral features of input spectra-pairs, are based on fully connected layers for efficient computing and share the parameters to ease the optimization. To solve the few samples problem, we propose a pseudo data generation method based on the linear mixed model and the assumption that background pixels are dominant in HSIs. For mitigating the impact of stochastic suboptimal initialization, we parallelly optimize several Siamese detectors with small computation burdens and aggregate them as ensembles in the inference time. The network ensembles outperform every detector in terms of stability and achieve an outstanding balance between background suppression and detection rate. Experiments on multiple data sets demonstrate that the proposed detector is superior to the state-of-the-art detectors.

Keywords: Siamese network; target detection; hyperspectral image; linear mixed model; ensemble method



Citation: Zhang, X.; Gao, K.; Wang, J.; Hu, Z.; Wang, H.; Wang, P. Siamese Network Ensembles for Hyperspectral Target Detection with Pseudo Data Generation. *Remote Sens.* **2022**, *14*, 1260. <https://doi.org/10.3390/rs14051260>

Academic Editor: Taejung Kim

Received: 18 February 2022

Accepted: 28 February 2022

Published: 4 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hyperspectral imaging is a developing area in remote sensing in which a hyperspectral spectrometer collects hundreds of narrow contiguous bands over a wide range of the electromagnetic spectrum [1]. Different from target detection in natural images [2], hyperspectral target detection aims to distinguish specific target pixels from the background in given HSIs with few prior spectral information of the target, which has been the focus of the remote sensing interpretation research [3].

In the past few decades, several classic hyperspectral target detectors have been proposed [3]. The spectral angular mapper (SAM) [4] and spectral information divergence (SID) [5] perform detections based on distance measurements. According to the background modeling, Zhang et al. [6] divided the traditional detection model into structured background models and unstructured background models. Structured background models include the constrained energy minimization (CEM) [7], the orthogonal subspace projection (OSP) [8], target-constrained interference minimize filter (TCIMF) [9], etc. Unstructured background models regard the background as samples from a multivariate Gaussian distribution, such as the generalized likelihood ratio test (GLRT) [10], adaptive coherence/cosine estimator (ACE) [11], and adaptive matched filter (MF) [12]. Most traditional detectors are

based on certain assumptions, which may fail in practice [13]. For instance, the background and target may not share the same spectral covariance matrix.

In the past few years, machine learning-based and sparse representation-based methods have improved detection performance. Hierarchical CEM detector (HCEM) [14], and ensemble-based constrained energy minimization (ECEM) [15], introduce hierarchical and ensemble structures to the CEM detectors, which improves the detection nonlinearity and generalization ability. The key idea of sparse representation-based methods is that each spectrum can be linearly combined with very few atoms of an over-complete dictionary. Typical methods include original sparse representation based target detector (STD) and its variants: sparse representation-based binary hypothesis detector (SRBBH) [16], coordinated-representation-based object detection (CR-TD) [17], and sparse and dense hybrid representation-based target detector (SDRD) [18]. Many scholars combined sparse representation and other techniques to develop many novel methods. Li et al. [17] proposed a combined sparse and collaborative representation (CSCR) for target detection, and Du et al. [19] proposed a hybrid sparsity and statistics-based method (HSSD). Zhao et al. [20] propose adaptive iterated shrinkage thresholding method (AISTM) for l_p -norm sparse representation and improve the detection performance.

Although traditional methods have achieved promising target detection results, there are still a few problems. For instance, the low spatial resolution of HSIs, atmospheric absorption, and scattering make the target spectra different from the prior spectra, introducing difficulties to the detection task [21]. Furthermore, the handcrafted spectral feature extraction filters restrict any further improvement of performance under complicated background interference [6]. Meanwhile, deep-learning-based methods have made great success in the remote sensing field, such as anomaly detection [22–25], classifications [26–28], image unmixing [29–32], etc. Many deep learning-based hyperspectral detectors [33–35] outperform the traditional methods relying on the feature extraction capability of neural networks. However, learning-based methods require numerous reliable training samples, but such samples are not directly obtainable from a single prior spectrum.

To overcome the few label samples problem, some unsupervised methods [35–38] exploit the neural networks as a feature extraction module and detect the target based on discriminative features with other detectors like CEM. Specifically, Shi et al. [36] introduce distance constraint stacked sparse autoencoders (SSAEs) to maximize the distinction between the target pixels and other background pixels and detect the target with a simple detector. Xie et al. [38] impose autoencoder and variational autoencoder for discriminative feature selection. Similarly, the authors of [37] select a subset of all bands for detection with a deep latent spectral representation learning-based autoencoder. In [35], a background-learning method is proposed and obtain the coarse detection map according to the reconstruction error. Although unsupervised-based methods improve the detection performance with discriminative features, these method requires other detectors for the final results in a two-stage way.

Many supervised methods have been proposed and integrate feature processing and detection. However, the unbalance between target and background samples is the critical problem of the supervised method. For generating numerous training samples, Zhang et al. [34] utilize generative adversarial networks to generate target and background spectra for pixels pairs. Gao et al. [21] generates simulated data with an auxiliary generative adversarial network. Zhu et al. [39] generate enough typical background pixels via a hybrid sparse representation and classification-based pixel selection strategy. Several convolutional network-based detectors have achieved great performance with numerous target and background spectra. Du et al. [40] propose a convolutional neural network-based target detector (CNNTD) and Zhang et al. [34] propose a novel target detection framework HTD-Net. HTD-Net and CNNTD feed the subtraction of pixels pairs to the multi-layers convolutional networks. Instead of feeding spectral subtraction to the networks. Zhu et al. [39] feeds the prior spectrum and the generated spectrum to a two-stream convolutional network (TSC-NTD) and detects the target with the subtraction of spectral features. Optimized with the

generated target and background spectra, supervised target detection methods [21,34,39,40] realized novel performance. However, convolutional-based feature extraction modules are designed with many convolutional layers, introducing much computation burden. In addition, the suboptimal initialization of these supervised methods will impact the performance of networks, and initializing networks with random parameters will lead to fluctuations in performance [41].

Siamese networks are capable of recognition with little available data with multiple networks sharing the parameters, proved by the scholars of [42], which is propitious to target detection with the single prior spectrum. This paper proposes a supervised Siamese fully connected target detector (SFCTD) composed of nonlinear feature extraction modules (NFEMs) and cosine angle distance-based classifiers. Two NFEMs, which extract discriminative spectral features of input spectra-pairs, are based on fully connected layers for efficient computing and share the parameters to ease the optimization. We utilize the cosine angle value of SAM measurement as the differential criterion to optimize the parameters of NFEMs. The cosine angle distances of spectral feature pairs represent the similarities of the input spectral pairs, serving as the target confidences of the test spectra. To solve the few samples problem, we propose a pseudo data generation method based on the linear mixed model and the assumption that background pixels are dominant in HSIs. For avoiding the impact of suboptimal initialization and achieve stable detection, we optimize several Siamese detectors independently and detect targets with the network ensembles.

The contributions of our work are summarized as follows.

- (1) A Siamese fully connected hyperspectral target detector (SFCTD) is proposed, consisting of nonlinear feature extraction modules and cosine angle distance based classifiers.
- (1) A pseudo data generation method is proposed to create numerous positive and negative spectral pairs with discrete similarity labels, i.e., 0 or 1. The SFCTD is effectively optimized with the generated spectral pairs.
- (3) A detection ensemble method is proposed for improving detection performance and stability. The Siamese detector ensembles outperform other state-of-the-art algorithms regarding the accuracy, recall, and background suppression, validated on multiple complex HSI data sets.

The remainder of this paper is organized as follows: Sections 2.2–2.4 introduce the methods of the proposed hyperspectral detector. Section 2.4 and 2.6 introduce the information of experimental data sets and implementation details. Section 3 presents the experimental results and ablation studies of the proposed method. The discussion and conclusion are drawn in Sections 4 and 5.

2. Materials and Methods

2.1. Abbreviations Define

For the convenience of the subsequent description, let $x_i \in \mathbb{R}^l$ denote the i -th spectral vector of the HSI $\mathbf{X} \in \mathbb{R}^{n \times l}$ with prior target spectrum $x^{prior} \in \mathbb{R}^l$, where n is the number of pixels in the HSI and l is the number of spectral bands. The generated training data set \mathbf{D} consists of positive spectral pairs, $\mathbf{Y}^+ = \{(y_i^+, y_i'^+) | i = 1, 2, \dots, n\}$ and negative spectral pairs $\mathbf{Y}^- = \{(y_i^-, y_i'^-) | i = 1, 2, \dots, n\}$, where y_i, y_i' represent spectral pairs associated with test spectrum x_i . The batch size of the training data set is denoted b and number of mini batches of an HSI with n pixels equals n/b . The NFEM is denoted f , and the test data pairs are denoted \mathbf{D}_{test} .

2.2. Siamese Hyperspectral Target Detector

As shown in Figure 1, the proposed Siamese detector consists of two NFEMs with shared structure and parameters. Each input spectral pair consists of a prior spectrum and a test spectrum. We separately feed the spectral pairs into each NFEM and compute the cosine angle distance of the transformed output features. The cosine angle values of SAM measurement represent the probabilities of the test spectra belonging to the target category.

To extract the features of the spectra effectively, i.e., 1-D vectors, the proposed NFEMs utilize fully connected networks instead of 1-D convolutional networks. Although convolutional networks have fewer parameters than fully connected networks because of the weight sharing of convolution operation, they have much more computation and random access memory burden. Specifically, each NFEM comprises a single batch norm layer and two fully connected blocks (FC blocks). Each FC block consists of a fully connected layer, a batch norm layer, and a nonlinear activation layer. We will illustrate each component successively in the following paragraph.

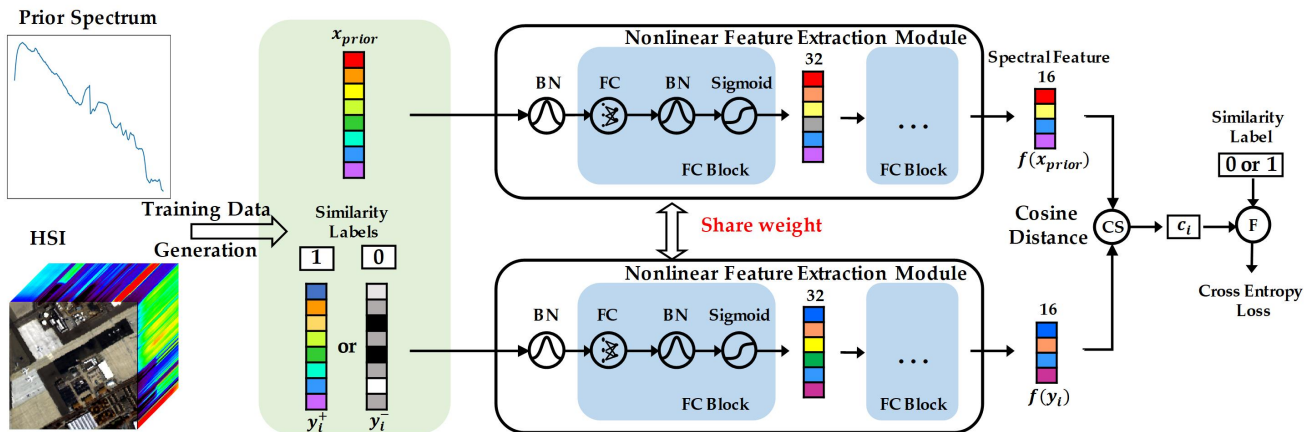


Figure 1. The training stage pipeline of the Siamese detector consisting of training data generation, spectral feature extraction, cosine distances and losses computation. x_{prior} , y_i^+ , and y_i^- represent the prior, target, and background spectrum, respectively. c_i represents c_i^- or c_i^+ , is the similarity label of the input spectral pair, (x_{prior}, y_i^-) or (x_{prior}, y_i^+) . Data generation provides spectral pairs with similarity labels. Each spectral pair is separately fed to weight-shared feature extraction modules. We optimize the network parameters with the cross-entropy loss between angle distance and similarity labels.

The amplitudes and waveforms of spectra in an HSI vary in positions because of different imaging and surface conditions, as shown in Figure 2a. Assuming that all the spectra belonging to one category are independently sampled from the same multidimensional random distribution, the distributions of target and background spectra are different because of their different physical property. However, the significant variance and mean shift of the test spectral distribution may impact the effectiveness of the feature extraction module. Hence, we preprocess the input spectral distribution at the beginning of feature extraction with a batch norm layer to reduce the shift and optimization difficulty. Instead of normalizing the spectral distribution with zero mean and unit variance, which may be unideal for optimizing loss functions, we use batch normalization (BN) to transform the original spectral distributions to distributions with learnable statistical parameters. Specifically, for the spectra mini batch $\mathbf{B}_m = \{(\mathbf{Y}_i^+, \mathbf{Y}_i^-) | i = b(m - 1), b(m - 1) + 1, \dots, bm\}$, $m \in \{1, 2, \dots, n/b\}$, BN normalizes the distribution of the spectra batch, \mathbf{B}_m , and transforms it to a distribution with a learnable mean β and variance γ^2 . The target spectra after BN is shown in Figure 2b. The equation of the BN process is:

$$\begin{cases} \mu_m = \frac{1}{4b} \sum_{k=b(m-1)}^{bm} \sum_{y \in \mathbf{Y}_k^+ \cup \mathbf{Y}_k^-} y \\ \sigma_m^2 = \frac{1}{4b} \sum_{k=b(m-1)}^{bm} \sum_{y \in \mathbf{Y}_k^+ \cup \mathbf{Y}_k^-} (y - \mu_m)^2 \\ BN(y) = \frac{(y - \mu_m) \times \gamma}{\sigma_m} + \beta, y \in \mathbf{Y}_i^+ \cup \mathbf{Y}_i^- \in \mathbf{B}_m \end{cases} \quad (1)$$

where μ_m, σ_m^2 are the mean and variance of input spectra and $\mathbf{Y}_i^+, \mathbf{Y}_i^-$ are the positive and negative spectral pairs for supervised learning optimization. The detailed method and purpose of generating these spectral pairs are illustrated in Section 2.3. In the training stage, μ_m and σ_m^2 change with the forward propagation of each mini-batch while β and γ^2 are optimized by the loss function in the backward times. In the testing stage, all the parameters are fixed for every process of each spectral pair. Experiments in Table 1 validate that *BN* enlarge the distribution difference between target and background spectra. We note that the learned parameters of the batch norm layers change with the HSIs and prior spectra.

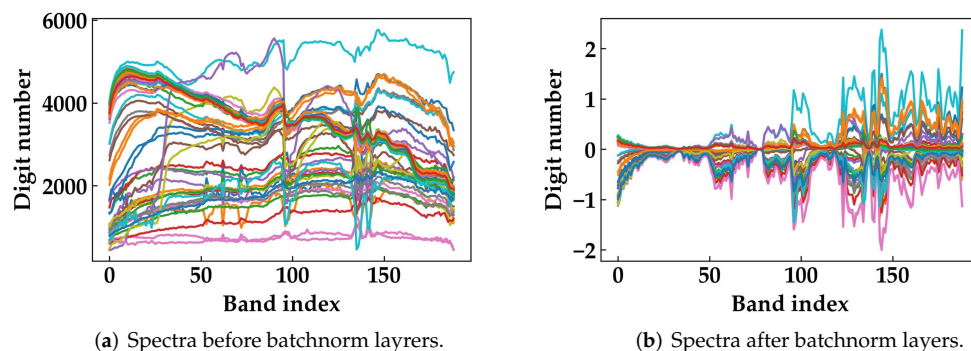


Figure 2. Visualization of the spectral distribution normalization. The batch norm layer mitigates the mean and variance shift of the input spectra-pairs. Each spectrum before and after the batch norm layers are painted with the same colors in (a,b). Different colors represent different spectral samples.

Table 1. Ablation studies of the batch norm layers and labeled data pairs creation method. The best results are reported in bold.

Data Sets	w/o BN	w/o Sigmod	w/o pos Pairs	All
San Diego Airport	0.8568	0.8678	0.9931	0.9941
Cuprite	0.4901	0.5917	0.8819	0.9656

After the *BN* operation, spectral pairs are separately fed to two FC blocks to generate discriminative spectral features. In each FC block, spectra are fed to the fully connected layer, batch norm layer, and nonlinear activation layer successively. A fully connected layer with weight $W = [w_1, w_2, \dots, w_{l'}], w \in \mathbb{R}^l$ transforms a spectrum with l band to a low-level feature space with l' dimension. Each vector $w_k, k \in [1, l']$ serves as a liner classifier for the spectra detection of the test HSI, which highlights background or target spectra to improve the spectral discriminability. Notably, the batch norm layers of the FC block play different roles to that of the preprocessing layers. The *BN* of the FC block converts the spectral features into the unsaturated interval of the activation function, which is usually operated before the nonlinear activation layers. The Sigmoid layer helps the NFEMs extract nonlinear spectral features for accurate detection. Finally, we obtain discriminative spectral features for cosine angle distance computation through spectra transformation of preprocessing and two FC blocks.

For the input spectral pair (y_i, y'_i) , the transformed spectral vector pair is $(f(y_i), f(y'_i))$. Different from [34,39], both of which utilize a fully connected layer to classify the feature subtraction of the input pairs, we derive a simple cosine angle distance classifier from SAM measurement. Specifically, we utilize the cosine angle distances of the two output vector pairs as the classification confidence, which equals the cosine angle values of SAM measurement. The angle distance is simple to compute and easy for derivation. The formula of the cosine angle distance-based classifier is as follows:

$$c_i = \frac{f(y_i) \cdot f(y'_i)}{\|f(y'_i)\|_2 \times \|f(y_i)\|_2} \tag{2}$$

where c_i is the cosine distance, $\|f(y'_i)\|_2$ is the Euclidean norm of vector $f(y'_i)$, and $|\cdot|$ represents the inner product of two vectors. Compared with subtraction of spectral pairs in [34] and spectral features subtraction in [39], the cosine angle distance of the proposed method is magnitude invariant.

The similarity label of the spectral pair is a discrete value, 0 or 1, which is the supervised label of the cosine distance. Considering the target detection as a classification problem, we utilized binary cross-entropy (BCE) to measure the distance between the similarity labels and cosine similarities. The optimization function L_m of mini batch \mathbf{B}_m is:

$$\begin{aligned} L_m &= \frac{1}{2b} \sum_{i=b(m-1)}^{bm} BCE(c_i^+, 1) + BCE(c_i^-, 0) \\ &= \frac{1}{2b} \sum_{i=b(m-1)}^{bm} -\log(c_i^+) - \log(1 - c_i^-) \end{aligned} \quad (3)$$

where c_i^+ denotes similarity of positive spectral pair, and c_i^- indicates the similarity of a negative spectral pair, and b represents the batch size. It is worth noting that the mini-batch \mathbf{B}_m includes b positive spectral pairs and b negative spectral pairs generated by the identical spectra of the HSI. A detailed description is illustrated in Algorithm 1.

Algorithm 1 Training stage of the SFCTD.

Input:

The detected HSI, $\{x_i | i = 1, 2, \dots, n\}$;
 Stochastic initialized Siamese detector $D_s, s \in \{1, 2, \dots, N\}$;
 Nonlinear feature extraction module of D_s, f_s ;
 The prior spectrum of the target, x^{prior} ;
 Batch size, b .

Generate labeled data pairs:

- 1: Generate negative spectral pair $\mathbf{Y}^- = \{\mathbf{Y}_i^- | i = 1, 2, \dots, n\}$ following Equation (4), where $\mathbf{Y}_i^- = (y_i^-, y_i^{-'})$;
- 2: Augment the target spectra following Equation (5) and generate positive spectral pairs $\mathbf{Y}^+ = \{\mathbf{Y}_i^+ | i = 1, 2, \dots, n\}$ following Equation (7), where $\mathbf{Y}_i^+ = (y_i^+, y_i^{+'})$;
- 3: Concatenate \mathbf{Y}^+ and \mathbf{Y}^- to obtain the training data set $\mathbf{D} = \{(\mathbf{Y}_i^+, \mathbf{Y}_i^-) | i = 1, 2, \dots, n\}$.

Forward and backward propagation of the Siamese detectors:

- 1: Shuffle the order of the spectral pairs of data set \mathbf{D} ;
 - 2: Feed each mini-batch $\mathbf{B}_m = \{(\mathbf{Y}_i^+, \mathbf{Y}_i^-) | i = b(m-1), b(m-1) + 1, \dots, bm\}$, $m \in \{1, 2, \dots, n/b\}$ to the Siamese network, obtaining transformed feature $f_s(\mathbf{B}_m)$;
 - 3: Compute the cosine angle distances of each transformed feature pairs following Equation (2), obtaining $\mathbf{C}_m = \{(c_i^+, c_i^-) | i = b(m-1), b(m-1) + 1, \dots, bm\}$;
 - 4: Compute the cross entropy distance between the detection results and labels with BCE follow Equation (3);
 - 5: Optimize the parameters of D_s with the BCE loss.
-

2.3. Pseudo Data Generation Method

We generate numerous pseudo data with positive and negative spectra-pairs to optimize the Siamese detectors. The negative spectral pairs comprise prior and background spectra, which are utilized to optimize the detector to filter the background pixels. In addition, the positive spectral pairs comprise prior and target spectra, which help to optimize the detector to distinguish target pixels with spectral variations. However, the known target and background spectra are not directly obtainable from the HSI. To solve this lack of data problem, we generate numerous background and target spectra based on the dominant background pixels and an LMM, as illustrated in Figure 3.

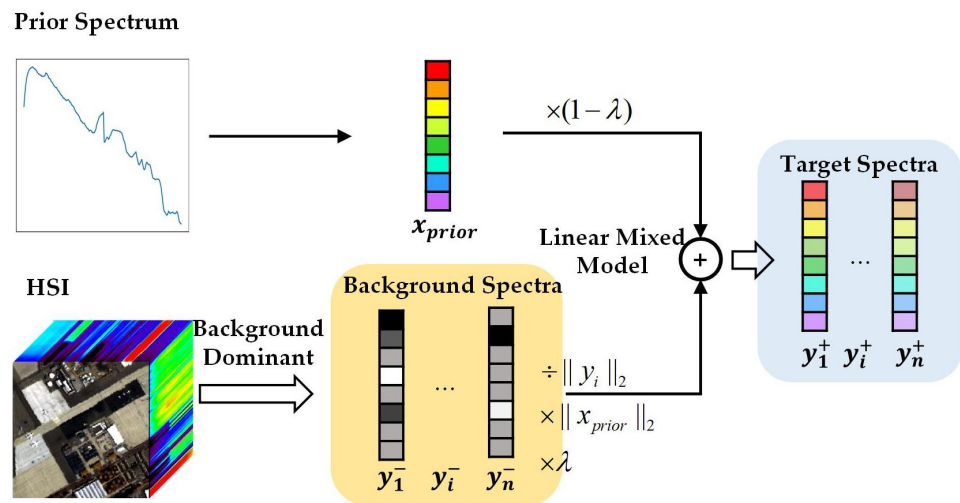


Figure 3. Training data generation method pipeline. y_i^-, y_i^+ are the background and simulated target spectrum, respectively. λ is the ratio of background spectra and equals 0.1.

To obtain background spectra from the test HSI, which contains both target and background spectra, we assume that background pixels are dominant in the HSIs. Based on this assumption, we consider each spectrum of the test HSI as a background spectrum and is different from the prior spectrum. The combination of background spectra and prior spectra make up negative spectral pairs Y_i^- , the specific formula of which is:

$$\begin{cases} Y_i^- = (y_i^-, y_i'^-) \\ y_i^- = x_i \\ y_i'^- = x^{prior} \end{cases}, i = 1, 2, \dots, n \quad (4)$$

Although a few target spectra may be labeled as background spectra mistakenly, this does not reduce the detection performance because the target pixels are far fewer than the correctly labeled background spectra. The effectiveness of the background spectra generation method is demonstrated by experiments conducted on several real data sets, as shown in Figure 4.

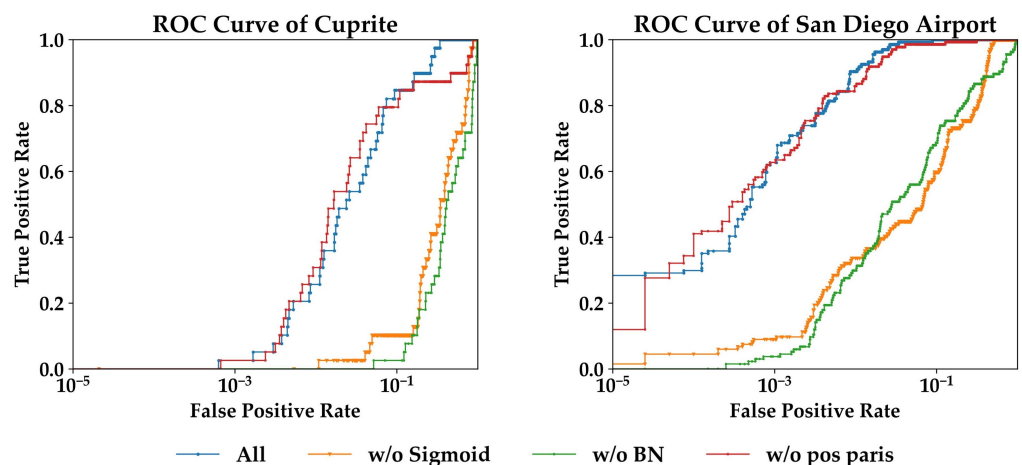


Figure 4. ROC curves of each ablation study for the San Diego Airport and Cuprite data sets.

To create multiple target spectra in addition to single prior spectra, we generate simulated target spectra by mixing up prior spectra and background spectra based on the LMM. The LMM assumes that the mixed spectrum x is a linear combination of target

spectrum e_t and background spectrum e_b with abundance coefficients a_t and a_b , respectively. The formula is as follows:

$$x = a_t \times e_t + a_b \times e_b \quad (5)$$

Since test spectra are different in amplitude, which may be much larger or smaller than the prior spectrum, we uniform the test spectra and adjust their amplitudes to those of the prior spectrum. The adjusted test spectra with small random weights multiplied are linearly mixed with the prior spectra, generating simulated target spectra, x_i^{mixed} . The visualization of background spectra and its associated simulated target spectra are exhibited in Figure 5, and the formula of the simulated target spectra generation is:

$$x_i^{mixed} = (1 - \lambda) \times x^{prior} + \lambda \times x_i \times \frac{\|x^{prior}\|_2}{\|x_i\|_2} \quad (6)$$

where λ is the ratio of background spectrum, and we set it as 0.1 for all the data sets. The abundance value of the target and background endmembers are 0.9 and 0.1, which means the resulting spectra is dominated by the target spectrum and can be seen as target spectra. It is worth noting that our target spectra generation method does not need to estimate the specific categories of the background spectra. Each spectrum e_b is regarded as the spectral noise added to the single prior spectrum. After obtaining the simulated target spectra, we combine the prior spectrum and target spectrum generating positive data pairs Y_i^+ , as follows:

$$\begin{cases} Y_i^+ = (y_i^+, y_i'^+) \\ y_i^+ = x_i^{mixed} \\ y_i'^+ = x^{prior} \end{cases}, i = 1, 2, \dots, n \quad (7)$$

The training data set, composed of positive and negative data pairs, is divided into $m = n/b$ mini-batches with batch size b . In each mini-batch, we use the identical spectra from the HSI to generate an equal number of positive and negative samples. Although the prior spectra in each mini-batch are the same, all the prior spectra are fed to the feature extraction module in the training stage for proper parameter updating of the batch norm layers.

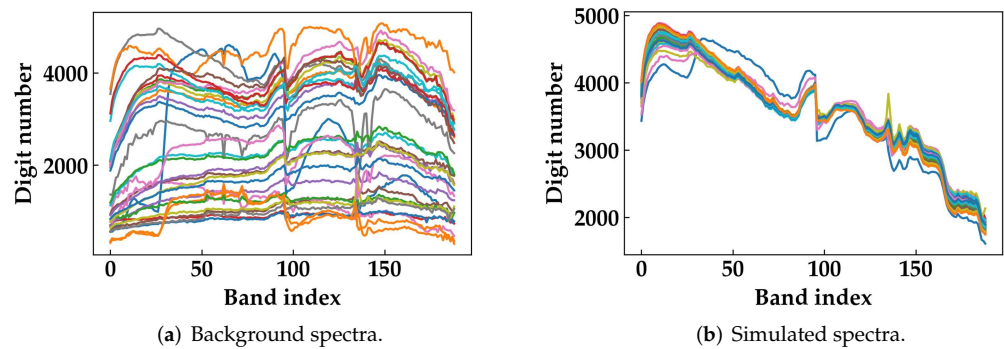


Figure 5. Visualization of simulated target spectra generation. Normalized background spectra are linearly mixed with prior spectra, generating augmented spectra. Each background spectra and associated simulated target spectra are painted with same colors in (a,b). Different colors represent different spectral samples.

2.4. Detection Ensemble Method

The performance of deep-learning-based detectors varies with model initialization, and suboptimal parameter initialization will impact the optimization and performance of the proposed detector. Specifically, detectors with stochastic initialization and data set shuffling will perform better than the average level. Although the probability of obtaining ideal parameter initialization is moderate, it is not easy to find the specific distribution of the ideal initialization. To achieve stable detection, we propose a simple but effective ensemble

method, as shown in Figure 6. Relying on the moderate probability of ideal stochastic initialization, we optimize and aggregate multiply Siamese detectors, D_1, \dots, D_N , to obtain a high-performance detector with a higher probability. Specifically, the final detection map \mathbf{C} is generated by averaging the detection map of each single Siamese detection following Equation (8). Experiments validate that the Siamese detector ensembles outperform every single detector, which means the multiple independently optimized detectors complement each other. As shown in Table 2, the ensemble result also shows better stability than each single detection result.

$$\mathbf{C} = \frac{1}{N} \sum_s^N \mathbf{C}_s \quad (8)$$

In the training stage, each Siamese detector is initialized with different stochastic parameters and trained with varying shuffles of the data sets, making sure multiple Siamese detectors are independent. Compared with other convolutional-based detectors [34,39,40], our proposed detector is based on the fully connected neural networks and the computing burden of which is much lower than that of convolutional-based detectors. Hence, we could parallelly optimize multiple detectors to improve detection stability using the parallel computing capability of GPU. In the testing stage, we follow the pipeline illustrated in Algorithm 2. Before the detection of N siamese detectors, we generate test spectra pairs by combining the prior spectrum x^{prior} and each test spectrum x_i . Then, spectral pairs of test data set \mathbf{Y}^{test} are fed to N Siamese detectors and generate detection maps, $\mathbf{C}_1, \dots, \mathbf{C}_N$. We ensemble all the results through the bagging approach to obtain the high-performance detection results without ground truth labels and manual intervention.

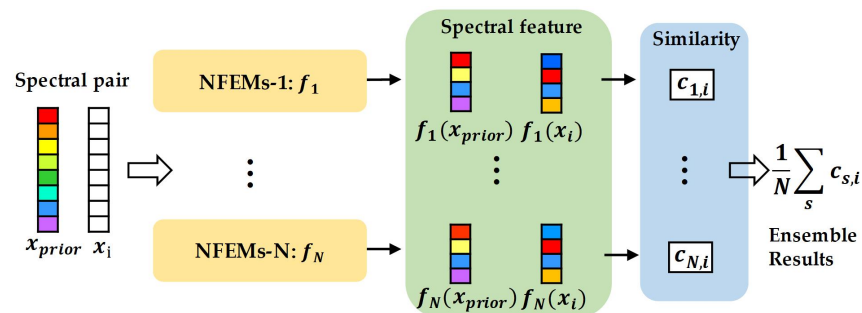


Figure 6. Pipeline of the detection ensemble method. x_i is the test spectrum to be detected, f_N represents the nonlinear feature extraction module (NFEM) of the N -th detector, and $c_{N,i}$ is the detection score of x_i outputted by the N -th detector.

Table 2. Ablation study results of the ensemble method. Four Siamese detectors ($N = 4$) are independently optimized and aggregated in the inference time. The best results are reported in bold.

Data	D-1		D-2		D-3		D-4		Ensembles	
	Mean	Std ($\times 10^{-2}$)	Mean	Std ($\times 10^{-2}$)	Mean	Std ($\times 10^{-2}$)	Mean	Std ($\times 10^{-2}$)	Mean	Std ($\times 10^{-2}$)
San Diego Airport	0.9915	0.261	0.9917	0.301	0.9936	0.202	0.9912	0.565	0.9941	0.183
San Diego Beach	0.9795	2.163	0.9874	0.827	0.9788	2.329	0.9889	0.802	0.9922	0.306
Texas Coast1	0.9909	0.713	0.9905	0.361	0.9923	0.272	0.9837	1.371	0.9938	0.239
Texas Coast2	0.9934	1.134	0.9965	0.252	0.9971	0.117	0.9973	0.205	0.9978	0.084
Synthetic (15 dB)	0.9948	0.261	0.9951	0.287	0.9952	0.235	0.9948	0.207	0.9966	0.179
Cuprite	0.9613	2.057	0.9566	2.148	0.9463	2.379	0.9516	1.978	0.9656	1.077

Algorithm 2 Test stage of the SFCTD.**Input:** input parametersThe detected HSI, $\mathbf{X} \in \mathbb{R}^{n \times l}$;The prior spectrum, x^{prior} ;Optimized N detectors, D_1, \dots, D_N .**Generate test data pairs:**

- 1: Duplicate the prior spectrum generating a prior spectra matrix, $\mathbf{X}^{prior} \in \mathbb{R}^{n \times l}$;
- 2: Generate test data pairs, $\mathbf{Y}^{test} = \{\mathbf{Y}_i^{test} | i = 1, 2, \dots, n\}$, where $\mathbf{Y}_i^{test} = (x_i, x^{prior})$.

Detection of N Siamese detectors.

- 1: Feed test spectral pairs \mathbf{Y}^{test} to the nonlinear feature extraction modules of each Siamese detector, obtaining transformed feature $f_s(\mathbf{Y}_i^{test})$, where $s \in \{1, 2, \dots, N\}$;
- 2: Compute the cosine similarity of the transformed features, obtaining each cosine similarity C_s , following Equation (2);
- 3: Average the similarity predictions of the N detectors to obtain the final results, C , following Equation (8).

2.5. Information of Experimental Data Sets

We used six real data sets and one synthetic data set to validate the proposed method, and the pseudo-color images are shown in the first row of Figure 7. All the real HSIs were captured by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). All the data sets we selected are used as the experimental data by other published hyperspectral target detection data sets.

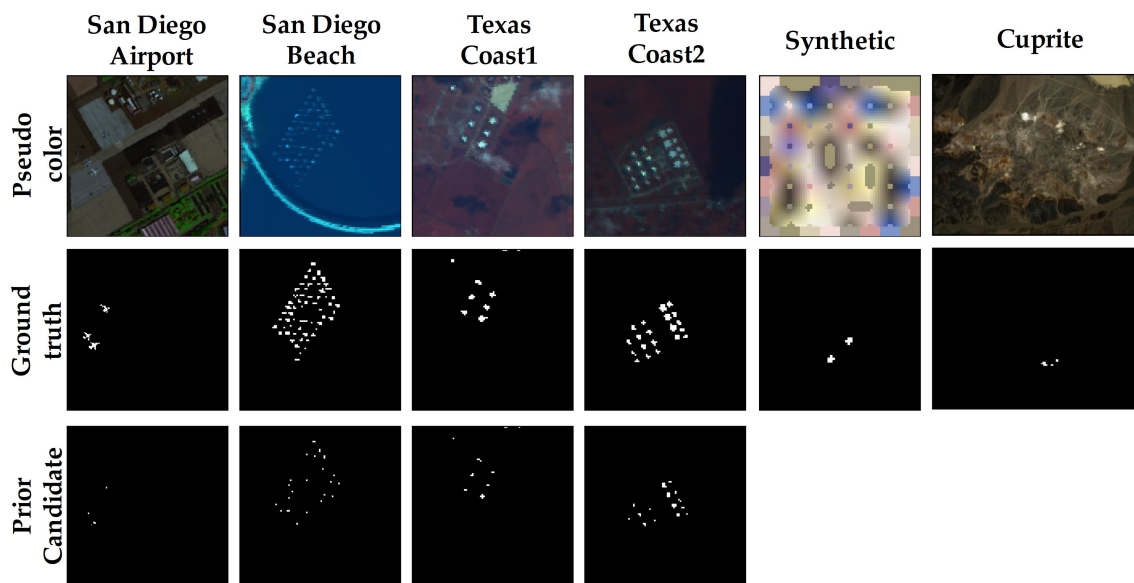


Figure 7. Six test data sets and ground truth maps. The first row shows the pseudo-color images of the test data sets. The second and third rows are the ground truth maps and prior spectra candidate maps. A 3×3 morphological erosion operation is conducted to the ground truth maps to obtain the prior spectra candidate maps. The Cuprite and Synthetic data sets have pure endmember spectra as priors.

All the data sets provide the target ground truth maps, but only two data sets (Cuprite and Synthetic) provide the prior spectra according to the USGS Digital Spectral Library [43]. For the data sets without prior spectra, scholars usually follow Manolakis et al. [3] to average the target spectra according to the ground truth maps as the prior simulated spectra. It is worth noting that the spectra of target boundary pixels reflect the interference of the target and background, which are different from the pure target spectra and are

impractical to obtain in actual application scenarios. Therefore, we did not average all the pixels of the ground truth map to generate the prior spectra.

Except for the Cuprite and Synthetic data sets, which provide pure endmember spectra as prior spectra, we conducted a morphological erosion operation to the ground truth maps to obtain the prior spectra candidate maps, and their average spectra represent the prior spectra. The Erosion operation is one of two basic operators in mathematical morphology. The erosion operation using B on binary ground truth map G is defined as:

$$G \ominus B = \{z \in E | B_z \subseteq G\} \quad (9)$$

where E is defined as Euclidean space and B_z is B shifted by z . This paper uses a positive 3×3 kernel as B . The morphological erosion is operated on the ground truth maps to discard the edge pixels. The ground truth annotation maps G , and the annotation maps after corrosion $G \ominus B$ operations are shown in the last two rows of Figure 7, respectively. The captured places and images details are listed below:

- **San Diego Airport:** The San Diego Airport data set was captured at San Diego with a 200×200 pixel size, which contains six airplanes and several backgrounds such as buildings and the parking apron. We selected three planes with the same size as targets, and target pixels annotation is the same as previous for a fair comparison, and the number of the target pixels is 134. Due to water-vapor absorption and atmospheric effects, we selected 189 bands from a total of 224 bands for the experiments.
- **San Diego Beach:** This hyperspectral data set was captured in San Diego with a spatial size 100×100 pixels. The scene of the data set is the beach, and the target pixels number is 202. The target annotation map refers to the data set from [21,38,44].
- **Texas Coast 1, Texas Coast 2:** These two hyperspectral data sets were captured along the Texan coast with a spatial size of 100×100 pixels. Several storage tanks were selected as targets from the urban scene, consisting of 67 and 155 pixels. The target annotation maps refer to the data set from [21,38,44]. The band number of the two data sets are 204 and 205.
- **Synthetic:** The spectra of the synthetic data set were generated from the USGS Digital Spectral Library, and there are 15 endmember spectra. For the comparison, we used the Labradorite HS17.3B endmember as the detection target, which is the same as [15]. Since all the methods used for comparison achieved 1 AUC in the clean data set, Gaussian white noise signal-to-noise ratios (SNRs) of 15 dB and 20dB were added to the original images.
- **Cuprite:** The Cuprite data set was captured in the Cuprite mining district of Nevada in 1997, where a subset of the images have a spatial size of 250×191 pixels. We selected buddingtonite from 14 kinds of minerals as the target for a fair comparison. The band number utilized is 188, and the spectra of buddingtonite in the United States Geological Survey (USGS) Digital Spectral Library was selected as a single prior spectrum.

2.6. Implementation Details

The experiments in the paper are run through Pytorch in Python on a computer with an Intel(R) Core(TM) i9-9900X CPU 3.50 GHz, GTX Titan Xp and 32G memory. We utilized the Adam optimizer for all the experiments and set the learning rate to 0.0005 and weight decay to 0.0005. Before training, the fully connected layers are initialized with a mean of 0 and a standard deviation of 0.001. We initialized the batch norm layers with β of 0 and γ of 1.

3. Results

In this section, we introduce the results of the experiments. First, we conduct the hyper-parameters sensitive experiments and select the proper parameters. Then, we conduct detection performance comparisons with seven state-of-the-art (SOTA) methods in terms of two-dimension receiver operating characteristic (2-D ROC) curves, area under the curve

(AUC) values. Since the 2-D ROC curve could not reflect the background suppression capabilities of the detectors, we supplement the box plots of detection confidences of all the comparison detectors for quantitative comparison and detection map visualizations for qualitative comparison. We use four Siamese detector ensembles for all the comparison experiments for consistency. The compared methods include four traditional methods (SAM, MF, ACE, and CEM), two advanced CEM-based methods (HCEM, ECEM), and one deep learning detector (TSCNTD) similar to our methods. Finally, we conduct an ablation experiment to study each component of the proposed method.

3.1. Hyper-Parameters Sensitive Experiments

Hyper-parameters are the parameters that are set before the training of networks. Ideal hyper-parameters could improve the performance of trained networks. Before comparing our method with other detectors, we first apply sensitive experiments of the hyper-parameters of the proposed method and select the proper hyper-parameters.

We set a different batch size, training epoch number, background spectra abundance λ , learning rate, and ensemble number and evaluate the detector performance under different hyper-parameters with AUC values in five test data sets except for the Cuprite data. For the setting of each parameter, we repeat ten times and compute the means and standard deviations of AUC values. The hyper-parameters candidates and experiment results are illustrated in Figure 8.

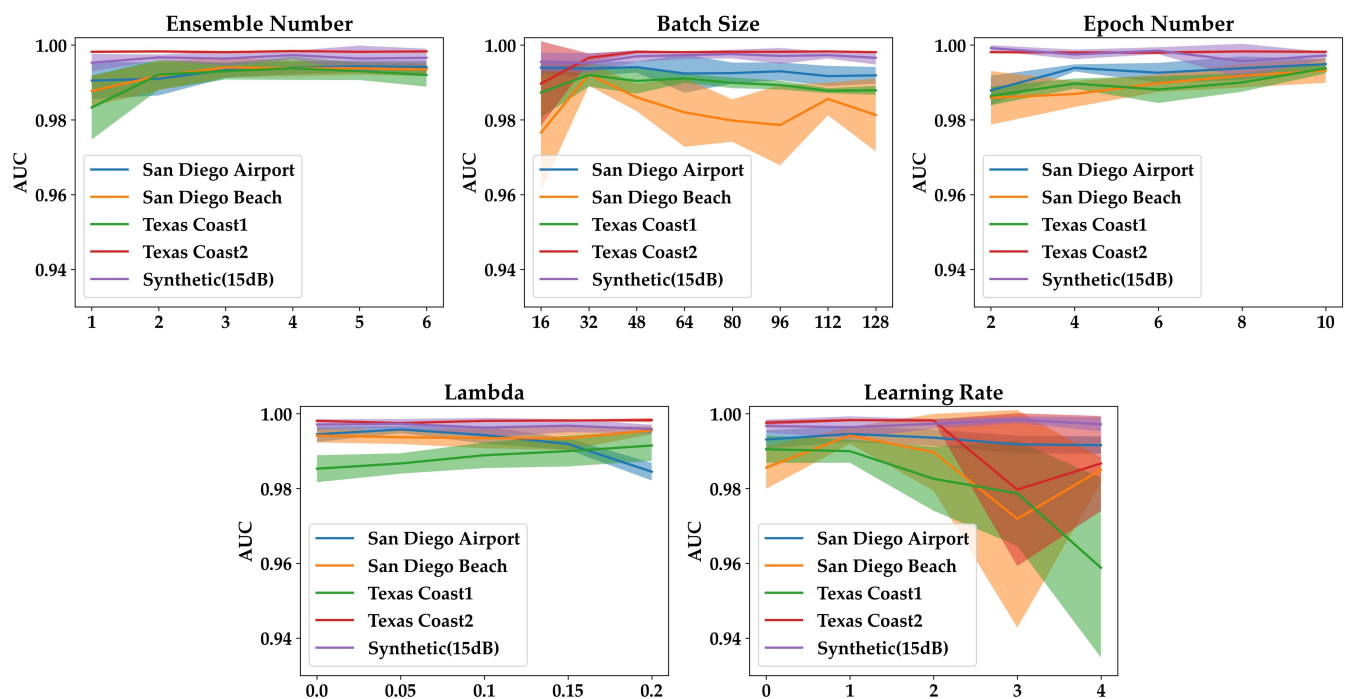


Figure 8. Parameters sensitivity experiments of the proposed Siamese detectors. The dark and light color plots represent the area under the curve (AUC) value means and standard deviation, respectively. The results of five data sets are painted with different colors. The default parameters, ensemble number = 4, batch size = 32, epoch number = 10, lambda = 0.1 and learning rate = 0.0005.

The experimental results can prove that background spectra abundance λ is not sensitive, and we select $\lambda = 0.1$ for all the data sets. The experimental results of epoch number and ensemble number reveal that the detection performance of the proposed Siamese detector is better and more stable with more training time and ensemble numbers. To make a trade-off between performance and efficiency, we optimize four Siamese detectors for ten epochs using the parallel computing capability of GPU. As for the learning rate, the

detector's performance with a learning rate larger than 5×10^{-3} is much more fluctuant. Therefore, we set the learning rate as 5×10^{-4} on all the data sets. Since the proposed detector is trained with fixed epoch numbers on all the data sets, the network may not be trained well with a large batch size because the iteration numbers are small. We set the batch size as 128 for the Cuprite data set because of its large image size and batch size 32 for all the other data sets.

3.2. Experiment Comparisons

3.2.1. Background Suppression Comparison

In Figure 9, six detection maps are visualized for the background suppression qualitative comparison. A higher visualization contrast-detection maps mean better background suppression capabilities. Among all the detectors, the ACE, HCEM, ECEM, TSCNTD, and proposed detector show higher visualization contrast than the SAM, MF, and CEM. However, the targets' integrity of our detection results is better than that of ACE, HCEM, and ECEM. Take, for example, the experiment on the San Diego Airport data set, the left and top planes of the ACE detection map are a bit blurred. The detection results of HCEM and ECEM fail to detect the margins of the target. Furthermore, the false alarm detection rate of the Siamese detector with NFEMs is better than that of TSCNTD; the latter detects many background pixels as the target.

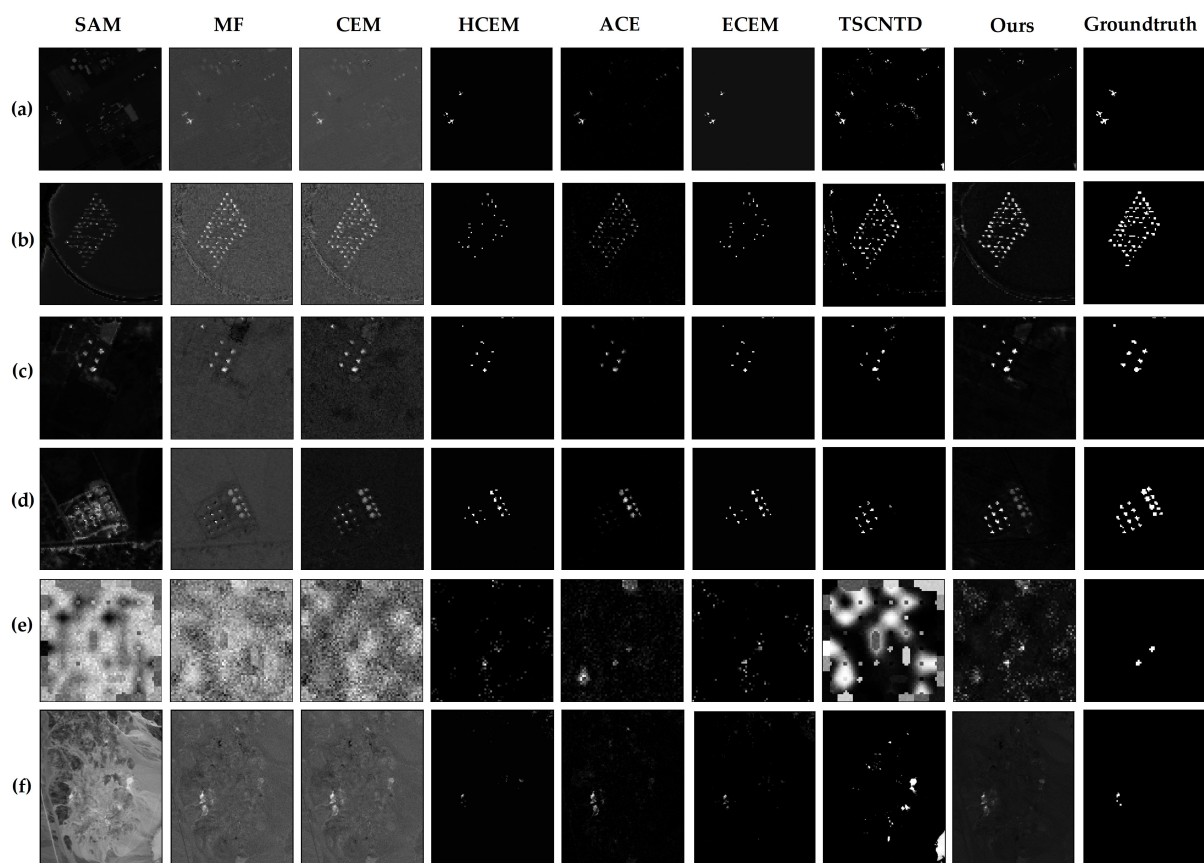


Figure 9. The visualization comparison of different algorithms. Exhibit detection maps of the (a) San Diego Airport, (b) San Diego Beach, (c) Texas Coast1, (d) Texas Coast2, (e) Synthetic (15 dB) and (f) Cuprite data sets.

Figure 10 shows a antitative comparison of background suppression. The red and green boxes reveal the confidence distributions of target and background pixels. Specifically, the wider a green box is, the larger the standard deviation of the background confidence distribution, which means background pixels' confidence is in an extensive range. The

lower a green box is, the smaller the mean of background confidence distribution is. Generally, the methods with good background suppression have flat and low green boxes and the red boxes whose lower quartiles are far away from the green boxes, which means most target pixels have higher confidence than background pixels. The box plot results prove the point consistent with that detection maps visualization results prove. CEM, SAM, and MF have wider and higher green boxes than the other methods. For the methods with low detection rates, such as HCEM, ECEM, and TSCNTD, the lower quartiles of their target boxes are close to zero in the data sets except for Synthetic, because many target pixels fail detect and their confidences are low. Although their green boxes are flat and low, the lower quartiles are close to the green boxes. For our method, the background boxes are flat and close to zero. Meanwhile, the lower quartiles of our target boxes are far away from the upper limits of the background boxes.

To sum up, our method achieves an outstanding balance between background suppression and target detection recall and outperforms the other comparison methods.

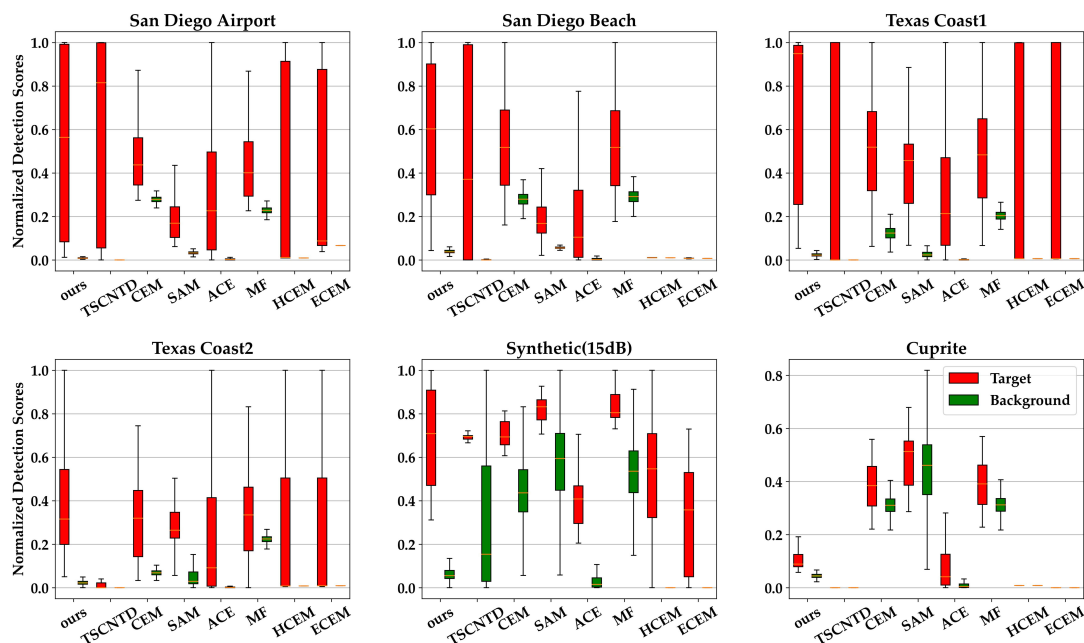


Figure 10. Boxplots of all the test algorithms on the six data sets. Each algorithm has two boxes on each data set. The red and green boxes represent the confidence distributions of target pixels and background pixels, respectively. The top of the boxes are the upper quartiles. The upper black lines and middle orange lines are the upper bounds and means of the distributions, respectively.

3.2.2. Roc Curves Comparison

ROC curves results are shown in Figure 11. ROC curve reflects the detection results in terms of detection rate and precision. The black line of each ROC curve represents the ensembled detection results of our method. The proposed Siamese detector shows a competitive detection rate in most false-positive rate thresholds for all the data sets. For the San Diego Airport data set, our method achieves a competitive detection rate under a low false-positive rate ($0-10^{-3}$) and the best detection rate as the growth of the false positive rate. Except for the San Diego Airport data set, our method surpasses all the comparison methods under almost all the false-positive rate thresholds. Especially for two data sets captured at Texas Coast, our method's curves are more than 20% higher than the other curves under 10^{-5} false-positive rate. For the San Diego Beach and Cuprite data sets, our curves outperform other methods almost 25% under the low false-positive rate between $10^{-3}-10^{-1}$.

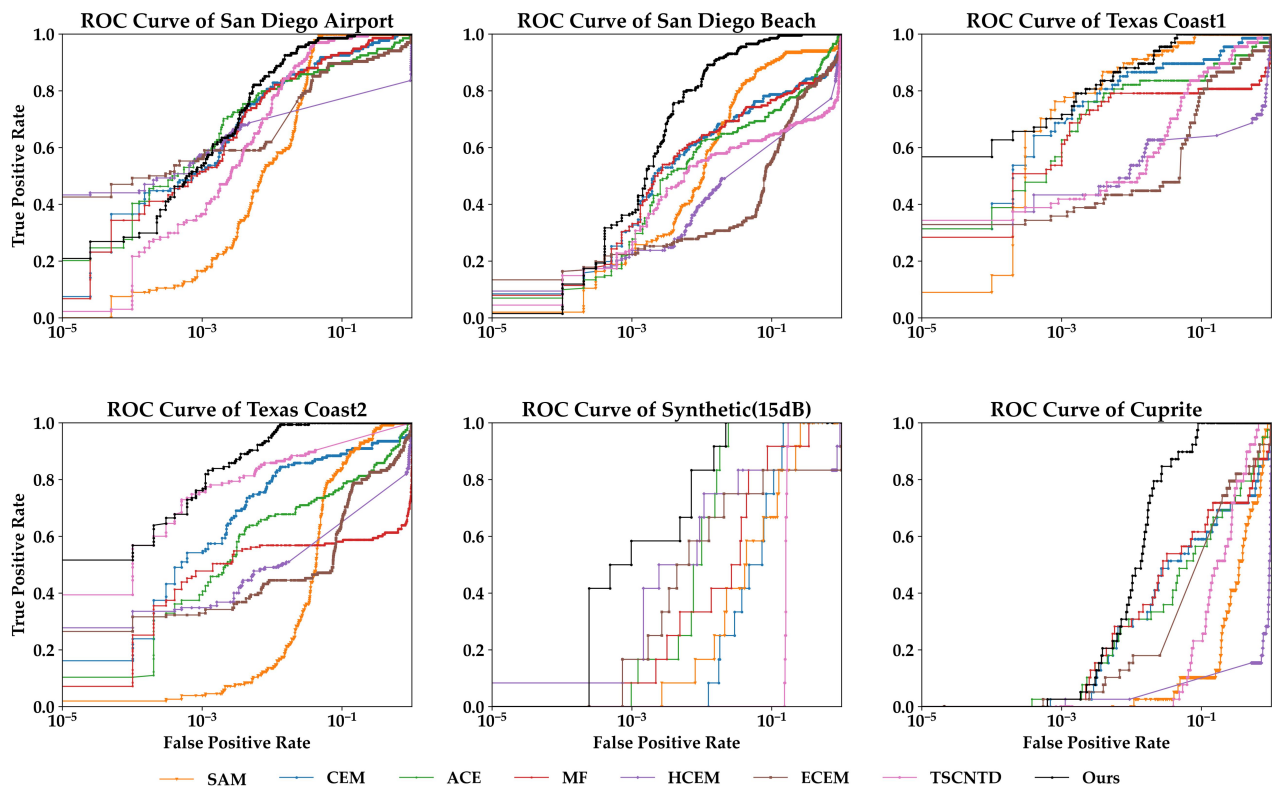


Figure 11. Receiver operating characteristic (ROC) curves of all the test algorithms on the six data sets. The black lines represent the proposed Siamese detector ensembles ($N = 4$).

3.2.3. AUC Values Comparison

Table 3 exhibits the AUC value results for all the test hyperspectral data sets except the Synthetic data set. The proposed Siamese network achieves the best AUC values in all the data sets except the Texas Coast 1 data set and surpasses the other methods by large margins, especially for the San Diego Beach, Cuprite, and Texas Coast 2 data sets. Specifically, our method outperforms the second-best methods by 0.008, 0.162, and 0.061 on the San Diego Beach, Cuprite, and Texas Coast 2 data sets. Since the Synthetic data set has random white noise, we repeated the test 10 times and calculated the mean and standard deviation of the AUC values, illustrated in Table 4. The proposed Siamese detector surpasses all the comparison methods under two noise conditions. The Siamese detector's lowest standard deviation of AUC values reflects its excellent detection stability under noise interference.

Table 3. The AUC values of different algorithms on five real data sets. The best results are reported in bold.

Method	San Diego Airport	San Diego Beach	Texas Coast 1	Texas Coast 2	Cuprite
SAM	0.9858	0.9276	0.9947	0.9368	0.5917
MF	0.9754	0.8449	0.8287	0.6219	0.7714
CEM	0.9719	0.8501	0.9628	0.9286	0.7480
HCEM	0.7637	0.7229	0.7123	0.7465	0.1737
ACE	0.9478	0.8641	0.9288	0.8851	0.7876
ECEM	0.9255	0.7912	0.8815	0.8288	0.8034
TSCNTD	0.9782	0.7217	0.8528	0.9156	0.7679
Ours	0.9941	0.9656	0.9938	0.9978	0.9922

Table 4. AUC values of the different algorithms applied to the Synthetic data set under the interference of white noise with two SNR conditions. The best results are reported in bold.

Method	Nosie of 15 dB SNR		Nosie of 20 dB SNR	
	Mean	Std	Mean	Std
SAM	0.9519	1.299×10^{-2}	0.9945	2.464×10^{-3}
MF	0.9630	1.824×10^{-2}	0.9891	5.383×10^{-3}
CEM	0.9488	1.222×10^{-2}	0.9815	4.805×10^{-3}
HCEM	0.9612	3.112×10^{-2}	0.9962	1.024×10^{-2}
ACE	0.9910	2.331×10^{-3}	0.9957	1.182×10^{-3}
ECEM	0.9336	6.142×10^{-2}	0.9718	3.749×10^{-2}
TSCNTD	0.8555	10.27×10^{-2}	0.8693	6.111×10^{-2}
Ours	0.9966	1.931×10^{-3}	0.9995	0.360×10^{-3}

3.3. Comparison with TSCNTD

To validate the superiority of our proposed Siamese detector to TSCNTD, we optimize the TSCNTD and Siamese detector with the same training data and compare the performance in terms of AUC values, test time, and stability. We repeat the training and testing ten times and compute the two methods' AUC value means and standard deviations. The experiment results are exhibited in Table 5.

According to the experimental results in Section 3.2 and AUC means of Table 5. The proposed Siamese detector outperforms the TSCNTD in detection recall and precision. As shown in the standard deviation results in Table 5, the Siamese detector ensembles outperform the TSCNTD in terms of stability with the help of the detection ensemble method. Moreover, by using a few fully connected layers rather than many convolutional layers, our method's test speed is six times faster than TSCNTD. Since the batch size and training epoch numbers are different, we only compare the test times for fair. In conclusion, the proposed Siamese detector ensembles are superior to the TSCNTD in performance, stability, and efficiency detection.

Table 5. Detection performance and efficiency comparison of the two-stream convolutional network-based target detector (TSCNTD) and our method. The best results are reported in bold.

Dataset	TSCNTD			Ours		
	AUC Mean	AUC Std	Test Time (ms)	AUC Mean	AUC Std	Test Time (ms)
San Diego Airport	0.9782	0.886×10^{-2}	43.88	0.9941	0.183×10^{-2}	8.975
San Diego Beach	0.7217	2.072×10^{-2}	7.978	0.9922	0.306×10^{-2}	0.9975
Texas Coast1	0.8528	9.639×10^{-2}	5.984	0.9938	0.239×10^{-2}	0.9973
Texas Coast2	0.9156	2.681×10^{-2}	5.984	0.9978	0.084×10^{-2}	0.9968
Synthetic (15 dB)	0.8568	4.599×10^{-2}	5.983	0.9966	0.233×10^{-2}	0.9971
Cuprite	0.7679	2×10^{-5}	400.8	0.9656	1.077×10^{-2}	4.985

3.4. Ablation Study

In this section, we study the effectiveness of batch norm layers, Sigmoid layers, generated positive spectra pairs, and detection ensemble methods. We don't present the detection performance optimized without negative spectral pairs because the positive data failed to optimize the network alone.

Figure 4 shows the 2-D ROC curves of the single Siamese detector without batch norm layers, Sigmoid layers, and positive spectra pairs compared with the normal one in two selected data sets. The ROC curves of the norm single Siamese detector are much better than that of the detector without Sigmoid layers and batch norm layers. Table 1 shows the AUC values of the curves in Figure 4. The Siamese detector with all the contributions has the largest AUC values.

To demonstrate the effectiveness of the proposed detection ensemble method. We optimize four randomly initialized Siamese detectors independently and compare their AUC values with ensembles. We repeat the experiment 10 times and exhibit the results in Table 2. The ensemble result surpasses all the other detectors in terms of standard

deviations and means of AUC values, which validates the stability improvement of the proposed detection ensemble method.

4. Discussion

There have a few supervised learning-based detectors similar to our method, CNNTD [40], HTD-Net [34] and TSCNTD [39]. These three methods adopt convolutional networks for spectral extraction and employ no nonlinear activation layers for the network structures. CNNTD and HTD-Net input the network with spectral differences, reducing feature discriminability. The TSCNTD cleverly designed two-stream networks that separately apply to the prior and other spectra and solve the problem in CNNTD and HTD-Net, which makes TSCNTD superior to CNNTD and HTD-Net [39]. However, TSCNTD uses two convolutional networks with nine layers to process the spectra pairs, making it slower than our method. Our proposed detector derives from the Siamese network, which is capable of recognition with little available data [42]. Similar to the network structures in [21,35,38], we design the Siamese detector with fully connected layers and nonlinear activation layers. Since Zhu et al. [39] has proved the superiority of TSCNTD to CNNTD and HTD-Net in terms of performance and speed, we only compare the proposed detector with TSCNTD.

Although TSCNTD has achieved great performance, the computation burden of convolutional networks is heavy. Moreover, the parameters are redundant and introduce optimization difficulty. Specifically, the upper stream is only responsible for the feature extraction of the prior spectrum, a constant vector, which makes up almost half the parameters. Hence, Zhu et al. [39] proposed a regularized cost to optimize these numerous parameters. Our method solves these problems with a Siamese detector that comprises two fully connected network sharing parameters. Parameters sharing reduces the number of parameters and reduces the difficulty of optimization. We also introduce nonlinear layers to improve the feature extraction capability. Experiments in Table 5 validate that the proposed detector is more effective than TSCNTD.

Shi et al. [41] propose a semisupervised domain adaptive few-shot learning (SDAFL) model and exhibit the standard deviation results to prove the detection stability of SDAFL. Other deep learning-based methods [34,39,40] do not give the standard deviation results. To study the detection stability of TSCNTD, we repeat the experiments of TSCNTD ten times and find the stability is unsatisfactory, as shown in Table 5. This paper pays attention to the detection stability and improves it with a classical machine learning method, ensemble learning. The detection ensemble method improves both the stability and performance but introduces computation.

For a Given HSI and prior spectrum, non-learning methods will give specific detection results, such as SAM, CEM, MF, ACE, HCEM, and ECEM. Our proposed Siamese detector outperforms in performance with the help of neural networks' excellent feature extraction capability. However, the parameters initialization of the neural networks introduces fluctuation in performance. Therefore, the repeatability of non-learning methods is better than TSCNTD and SFCTD.

5. Conclusions

This paper proposes a Siamese fully connected network-based hyperspectral target detector, denoted as SFCTD, consisting of two nonlinear feature extraction modules (NFEMs) and a cosine angle distance-based classifier. Two Siamese structured NFEMs share the parameters and extract the discriminative features of prior and test spectra, respectively. The cosine angle distances of spectral feature pairs measure the confidence of test spectra regarded as the target. The SFCTD is effectively optimized by the generated pseudo positive and negative spectral pairs. To mitigate the performance fluctuation caused by the random initialization of parameters, we parallelly optimize several SFCTD, and the network ensembles are more stable than single SFCTD. Experiment results validate that the proposed SFCTD outperforms non-learning detectors in performance, such as SAM, MF,

CEM. In speed and performance, the SFCTD is superior to the similar supervised learning method, TSCNTD. The SFCTD has a lightweight structure and achieves an excellent trade-off between detection accuracy and computational cost, suitable for conditions with insufficient preliminary data. In the future, we will investigate the SFCTD's capability to detect the same targets in similar HSIs with different imaging conditions.

Author Contributions: X.Z. and J.W. conceived and designed the study. X.Z. constructed the model, implemented the experiments, and drafted the manuscript. Z.H. and H.W. contributed to improving the manuscript, and P.W. collected the hyperspectral data sets. K.G. provided the overall guidance to this work and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology (Grant No. GZZKFJJ2020004), the National Natural Science Foundation of China (Grant Nos. 61875013 and 61827814), and the Natural Science Foundation of Beijing Municipality (Grant No. Z19J00064).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. The datasets can be found here: https://github.com/Rui-ZHAO-ipc/E_CEM-for-Hyperspectral-Target-Detection, (accessed on 18 February 2022) and <http://xudongkang.weebly.com/data-sets>, accessed on 18 February 2022.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nasrabadi, N.M. Hyperspectral target detection: An overview of current and future challenges. *IEEE Signal Process. Mag.* **2013**, *31*, 34–44. [[CrossRef](#)]
2. Andriyanov, N.A.; Dementiev, V.E.; Tashlinskii, A.G. Detection of objects in the images: From likelihood relationships towards scalable and efficient neural networks. *Comput. Opt.* **2022**, *46*, 139–159.
3. Manolakis, D.; Marden, D.; Shaw, G.A. Hyperspectral image processing for automatic target detection applications. *Linc. Lab. J.* **2003**, *14*, 79–116.
4. Jin, X.; Paswaters, S.; Cline, H. A comparative study of target detection algorithms for hyperspectral imagery. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV, Proceedings of the SPIE Defense, Security, and Sensing, Orlando, FL, USA, 13–17 April 2009*; SPIE: Bellingham, WA, USA, 2009; Volume 7334, p. 73341W.
5. Chang, C.I. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Trans. Inf. Theory* **2000**, *46*, 1927–1932. [[CrossRef](#)]
6. Zhang, L. Advance and future challenges in hyperspectral target detection. *Geomat. Inf. Sci. Wuhan Univ.* **2014**, *39*, 1387–1394.
7. Farrand, W.H.; Harsanyi, J.C. Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique. *Remote Sens. Environ.* **1997**, *59*, 64–76. [[CrossRef](#)]
8. Harsanyi, J.C.; Chang, C.I. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 779–785. [[CrossRef](#)]
9. Ren, H.; Chang, C.I. Target-constrained interference-minimized approach to subpixel target detection for hyperspectral images. *Opt. Eng.* **2000**, *39*, 3138–3145. [[CrossRef](#)]
10. Kraut, S.; Scharf, L.L.; McWhorter, L.T. Adaptive subspace detectors. *IEEE Trans. Signal Process.* **2001**, *49*, 1–16. [[CrossRef](#)]
11. Kraut, S.; Scharf, L.L.; Butler, R.W. The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic. *IEEE Trans. Signal Process.* **2005**, *53*, 427–438. [[CrossRef](#)]
12. Robey, F.C.; Fuhrmann, D.R.; Kelly, E.J.; Nitzberg, R. A CFAR adaptive matched filter detector. *IEEE Trans. Aerosp. Electron. Syst.* **1992**, *28*, 208–216. [[CrossRef](#)]
13. Sakla, W.; Chan, A.; Ji, J.; Sakla, A. An SVDD-based algorithm for target detection in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2010**, *8*, 384–388. [[CrossRef](#)]
14. Zou, Z.; Shi, Z. Hierarchical suppression method for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 330–342. [[CrossRef](#)]
15. Zhao, R.; Shi, Z.; Zou, Z.; Zhang, Z. Ensemble-based cascaded constrained energy minimization for hyperspectral target detection. *Remote Sens.* **2019**, *11*, 1310. [[CrossRef](#)]
16. Zhang, Y.; Du, B.; Zhang, L. A sparse representation-based binary hypothesis model for target detection in hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *53*, 1346–1354. [[CrossRef](#)]

17. Li, W.; Du, Q.; Zhang, B. Combined sparse and collaborative representation for hyperspectral target detection. *Pattern Recognit.* **2015**, *48*, 3904–3916. [[CrossRef](#)]
18. Guo, T.; Luo, F.; Zhang, L.; Tan, X.; Liu, J.; Zhou, X. Target detection in hyperspectral imagery via sparse and dense hybrid representation. *IEEE Geosci. Remote. Sens. Lett.* **2019**, *17*, 716–720. [[CrossRef](#)]
19. Du, B.; Zhang, Y.; Zhang, L.; Tao, D. Beyond the sparsity-based target detector: A hybrid sparsity and statistics-based detector for hyperspectral images. *IEEE Trans. Image Process.* **2016**, *25*, 5345–5357. [[CrossRef](#)]
20. Zhao, X.; Li, W.; Zhang, M.; Tao, R.; Ma, P. Adaptive iterated shrinkage thresholding-based lp-norm sparse representation for hyperspectral imagery target detection. *Remote Sens.* **2020**, *12*, 3991. [[CrossRef](#)]
21. Gao, Y.; Feng, Y.; Yu, X. Hyperspectral Target Detection with an Auxiliary Generative Adversarial Network. *Remote Sens.* **2021**, *13*, 4454. [[CrossRef](#)]
22. Ma, N.; Yu, X.; Peng, Y.; Wang, S. A lightweight hyperspectral image anomaly detector for real-time mission. *Remote Sens.* **2019**, *11*, 1622. [[CrossRef](#)]
23. Ran, Q.; Liu, Z.; Sun, X.; Sun, X.; Zhang, B.; Guo, Q.; Wang, J. Anomaly Detection for Hyperspectral Images Based on Improved Low-Rank and Sparse Representation and Joint Gaussian Mixture Distribution. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6339–6352. [[CrossRef](#)]
24. Xie, W.; Zhang, X.; Li, Y.; Lei, J.; Du, Q. Weakly Supervised Low-Rank Representation for Hyperspectral Anomaly Detection. *IEEE Trans. Cybern.* **2021**, *51*, 3889–3900. [[CrossRef](#)] [[PubMed](#)]
25. Fu, X.; Jia, S.; Zhuang, L.; Xu, M.; Zhou, J.; Li, Q. Hyperspectral Anomaly Detection via Deep Plug-and-Play Denoising CNN Regularization. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 9553–9568. [[CrossRef](#)]
26. Fang, B.; Bai, Y.; Li, Y. Combining spectral unmixing and 3d/2d dense networks with early-exiting strategy for hyperspectral image classification. *Remote Sens.* **2020**, *12*, 779. [[CrossRef](#)]
27. Xi, B.; Li, J.; Li, Y.; Song, R.; Sun, W.; Du, Q. Multiscale context-aware ensemble deep KELM for efficient hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 5114–5130. [[CrossRef](#)]
28. Shen, Y.; Zhu, S.; Chen, C.; Du, Q.; Xiao, L.; Chen, J.; Pan, D. Efficient deep learning of nonlocal features for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 6029–6043. [[CrossRef](#)]
29. Ahmed, A.M.; Duran, O.; Zweiri, Y.; Smith, M. Hybrid spectral unmixing: Using artificial neural networks for linear/non-linear switching. *Remote Sens.* **2017**, *9*, 775. [[CrossRef](#)]
30. Dou, Z.; Gao, K.; Zhang, X.; Wang, H.; Wang, J. Hyperspectral unmixing using orthogonal sparse prior-based autoencoder with hyper-Laplacian loss and data-driven outlier detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6550–6564. [[CrossRef](#)]
31. Nalepa, J.; Myller, M.; Tulczyjew, L.; Kawulok, M. Deep Ensembles for Hyperspectral Image Data Classification and Unmixing. *Remote Sens.* **2021**, *13*, 4133. [[CrossRef](#)]
32. Li, H.; Feng, R.; Wang, L.; Zhong, Y.; Zhang, L. Superpixel-based reweighted low-rank and total variation sparse unmixing for hyperspectral remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 629–647. [[CrossRef](#)]
33. Li, W.; Wu, G.; Du, Q. Transferred deep learning for hyperspectral target detection. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 5177–5180.
34. Zhang, G.; Zhao, S.; Li, W.; Du, Q.; Ran, Q.; Tao, R. HTD-net: A deep convolutional neural network for target detection in hyperspectral imagery. *Remote Sens.* **2020**, *12*, 1489. [[CrossRef](#)]
35. Xie, W.; Zhang, X.; Li, Y.; Wang, K.; Du, Q. Background learning based on target suppression constraint for hyperspectral target detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5887–5897. [[CrossRef](#)]
36. Shi, Y.; Lei, J.; Yin, Y.; Cao, K.; Li, Y.; Chang, C.I. Discriminative feature learning with distance constrained stacked sparse autoencoder for hyperspectral target detection. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1462–1466. [[CrossRef](#)]
37. Xie, W.; Lei, J.; Yang, J.; Li, Y.; Du, Q.; Li, Z. Deep latent spectral representation learning-based hyperspectral band selection for target detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 2015–2026. [[CrossRef](#)]
38. Xie, W.; Yang, J.; Lei, J.; Li, Y.; Du, Q.; He, G. SRUN: Spectral regularized unsupervised networks for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 1463–1474. [[CrossRef](#)]
39. Zhu, D.; Du, B.; Zhang, L. Two-Stream Convolutional Networks for Hyperspectral Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 6907–6921. [[CrossRef](#)]
40. Du, J.; Li, Z.; Sun, H. CNN-based target detection in hyperspectral imagery. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 2761–2764.
41. Shi, Y.; Li, J.; Li, Y.; Du, Q. Sensor-independent hyperspectral target detection with semisupervised domain adaptive few-shot learning. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 6894–6906. [[CrossRef](#)]
42. Koch, G.; Zemel, R.; Salakhutdinov, R. Siamese neural networks for one-shot image recognition. In Proceedings of the ICML Deep Learning Workshop, Lille, France, 6–11 July 2015; Volume 2.
43. Clark, R.N.; Swayze, G.A.; Gallagher, A.J.; King, T.V.; Calvin, W.M. *The US Geological Survey, Digital Spectral Library: Version 1 (0.2 to 3.0 μm)*; Technical Report; Geological Survey (US): Reston, VI, USA, 1993.
44. Kang, X.; Zhang, X.; Li, S.; Li, K.; Li, J.; Benediktsson, J.A. Hyperspectral anomaly detection with attribute and edge-preserving filters. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5600–5611. [[CrossRef](#)]