

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

Weighted Discriminative Collaborative Competitive Representation with Global Dictionary for Hyperspectral Target Detection

Rong Liu, Jiake Wu, Dehui Zhu, and Bo Du, *Senior Member, IEEE*

Abstract—Hyperspectral target detection is a promising yet challenging endeavor in remote sensing image processing. Representation learning-based detectors have become one of the mainstream methods to address the task. However, these methods often suffer from weak separability between the target and the background, which results in inferior target detection performance. The reason relies on that their detection model cannot effectively distinguish the subtle differences between the target and the background. To tackle this issue, this paper proposes a new weighted discriminative collaborative competitive representation model (denoted as WDCCR) for hyperspectral target detection. In WDCCR, the separability between targets and backgrounds is enhanced by integrating discriminative, competitive, and weight constraints. Meanwhile, to obtain pure background pixels for the representation model, we investigate a new category-based pixel selection method with target orthogonal purification. The proposed WDCCR target detector is evaluated on six hyperspectral datasets. Experimental results demonstrate that WDCCR outperforms other advanced methods, achieving good detection performance in hyperspectral target detection. The code will be available at <https://github.com/liurongwhm>.

Index Terms—Hyperspectral remote sensing, collaborative competitive representation, dictionary learning, target detection.

I. INTRODUCTION

Hyperspectral images (HSIs) stand out for their capacity to offer rich spectral information through hundreds of approximately continuous spectral bands [1-2]. With its fine spectral resolution, HSI is particularly effective in discriminating substances even with minimal spectral differences [3-4], making it indispensable across various applications [5-6]. Notably, hyperspectral target detection (HTD) emerges as an area of considerable promise, leveraging the capabilities of HSI to identify camouflaged or complex background targets, which is crucial in mineral exploration, military reconnaissance, and environmental monitoring [7-8].

This work was supported by the National Natural Science Foundation of China under Grant 62201622 and the China National Postdoctoral Program for Innovative Talents under Grant BX20240492. (*Corresponding author: Dehui Zhu*).

Rong Liu and Jiake Wu are with the School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China (e-mail: liurong25@mail.sysu.edu.cn; wujk8@mail2.sysu.edu.cn).

Dehui Zhu is with the National Key Laboratory of Automatic Target Recognition, College of Electrical Science and Technology, National University of Defense Technology, Changsha 410073, China (e-mail: dhzhu95@163.com).

Bo Du is with the School of Computer Science, National Engineering Research Center for Multimedia Software, Institute of Artificial Intelligence, Wuhan University, Wuhan 430072, China (e-mail: gunspace@163.com).

When the prior target spectrum is available, HTD can essentially be considered as a binary classification problem, where test pixels are categorized as target or background. However, the methods devised for classification prove inapplicable for target detection for two reasons. First, only a few target pixels cannot accurately estimate the statistical properties of the target class from training data and the evaluation criteria for classification are therefore not applicable. Besides, the limited spatial resolution of HSIs means that certain targets may only be detectable at low probabilities or as part of a single pixel or subpixel [9]. Therefore, the framework for the design of target detection is usually based on the Neyman-Pearson criterion, which maximizes the probability of detection under a low false alarm rate [10].

Over the past three decades, numerous algorithms have been developed for HTD. These HTD algorithms can be roughly divided into three main categories.

(1) Classical detection algorithms: HTD detectors rely on spectral differences between the target and other land covers as the basis for detection at an early stage. This involves using techniques like Euclidean distance (ED) [11] or spectral angle mapping (SAM) [12] to compare test pixels with known spectra. The advent of statistical methods in signal processing leads to the proposal of statistical-based approaches. Examples include the adaptive coherence estimator (ACE) [13] and the matched filter (MF) [14], which treated target detection as a binary hypothesis test, relying on Gaussian distribution assumptions. The constrained energy minimization (CEM) [15] method designs a finite-impulse filter to highlight targets while suppressing background signals. Subsequently, variants of MF, such as adaptive matched filter (AMF) [16], were developed. Alternatively, some methods project hyperspectral signals into subspace models to heighten target-background contrast, such as the orthogonal subspace projection (OSP) [17] algorithm, improving signal-to-noise ratio by projecting image data into a subspace orthogonal to the background. The detection performance of these classical detection methods is often constrained by the requirement of certain statistical assumptions. However, the dependence on particular statistical assumptions often fails to correspond accurately with the complexities of real-world hyperspectral images, which results in limited target detection performance.

(2) Machine learning-based algorithms: These methods are developed using different machine learning techniques and

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

widely applied in different HSI processing tasks [18,19]. For example, kernel-based methods are proposed based on kernel technology, projecting data into higher kernel space to achieve linearly separable. Typical detectors are the kernel constrained energy minimization (KCEM) [20], the kernel spectral matched filter (KSMF) [21], and the kernel orthogonal subspace projection (KOSP) [22]. Other popular machine learning-based methods are representation learning-based detectors, which are developed from pattern recognition theory, and gain prominence in target detection, as they do not require the estimation of specific statistical assumptions. In 2011, sparse representation (SR) was introduced for RS image classification and target detection and has raised a stream to study the sparsity of RS data in subsequent years [23]. Sparse representation target detectors (STD) [24], sparse representation and binary hypothesis (SRBBH) [25], and combined sparse and collaborative representation (CSCR) [26] represent each test pixel as a linear combination of atoms from the background dictionary or target dictionary and then detects targets by using representation residuals. Spatially adaptive sparse representation for target detection (SASTD) [27] and weighted Cauchy distance graph and local adaptive collaborative representation (WCDG) [28] incorporate spatial information into the representation. Collaborative-guided spectral abundance learning model (CGSAL) [29] applies the collaborative constraint in the bilinear mixing model. Given the potential for background contamination by the sliding dual window in the aforementioned methods, several strategies have been developed to obtain a pure background dictionary, as explored in [30-34]. Other machine learning-based algorithms have garnered considerable attention as well. To tackle the challenge of distinguishing target pixels from background pixels, particularly when they share similar spectra due to subpixel targets, Zhang et al. [35] introduce metric learning. The approach aims to enhance discrimination between target spectra by maximizing the distance between positive and negative samples. Concurrently, Dong et al. [36] delve into multi-metric learning utilizing random forests. In [37] and [38], researchers propose multiple-instance learning frameworks to confront the issue of training samples with imprecise labels. These frameworks leverage datasets labeled as positive and negative bags, thereby addressing the challenge posed by ambiguous labels. According to the low-rank property of the natural image [39], some scholars pay attention to take use of spatial and temporal information based on tensor decomposition [40]. Moreover, some HSI preprocessing methods are specifically designed for target detection, including band selection [41-42], noise reduction [43], and band fusion [44].

(3) Deep learning-based detectors: Deep learning-based methods have experienced a surge in the field of HSIs, attributed to their remarkable feature extraction capabilities [45-46]. The integration of deep neural networks into HTD frameworks has become a prevalent approach [47-49]. For example, the hyperspectral target detection network (HTD-Net)

utilizes pixel pairs to train a similarity discrimination model [48]. Some scholars have leveraged autoencoders to extract features or bands from images for subsequent detection. Moreover, HTD-Net incorporates spatial features into its detection process. For instance, Xie et al. [50] propose a detector incorporating a spectral generative adversarial network, while Qin et al. [51] introduce a vision transformer for learning discriminative spectral-spatial features. Others, like Yand et al. [52], emphasize extracting both spatial and spectral information while preserving local spatial structure. Additionally, recent advancements include novel approaches such as the hyperspectral target detector based on an interpretable representation network (HTD-IRN), which employs a subspace representation network to realize the linear mixture model (LMM) for HTD [53]. Zhao et al. [54] integrate object detection methods from computer vision and contextual edge map techniques to first locate subregions of targets. Deep learning-based methods require accurate prior knowledge to train the network. However, the target detection task often lacks sufficient samples, making it challenging to construct appropriate training datasets. The need for abundant training samples and the significant time investment for network training may render these methods suboptimal. Furthermore, many data-driven deep learning-based algorithms struggle to provide a clear physical interpretation of the learned features in remote sensing data and lack a theoretical explanation for the relationship between input observations and output predictions [23].

Although many scholars have endeavored to address HTD by proposing a variety of methodologies, a prevalent issue with most extant models lies in their deficiencies of the detection model to effectively discriminate the target and diverse backgrounds due to the nuanced spectral difference. This inadequacy stems from the failure to express enough otherness between the target and backgrounds adaptively, resulting in weak separability and inferior target detection performance.

To address this issue, this study investigates a novel representation learning model to enhance separability between the target and the background, termed the weighted discriminative collaborative competitive representation model, denoted as WDCCR. The proposed WDCCR improves the separability by integrating three specific constraints, i.e., discriminative, competitive, and weight constraints. Besides, global background dictionary construction and target pixel augmentation are conducted to provide sufficient samples for the representation model.

The main contributions of this work are as follows:

(1) A new discriminative representation learning model is investigated in this paper, in which discriminative, competitive, and weight constraints are imposed to effectively integrate inter-class discrimination and enhance the separability of the detection model. Specifically, the competitive constraint distinguishes the reconstructive residuals on the basis of overall residual, while the discriminative constraint decreases

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

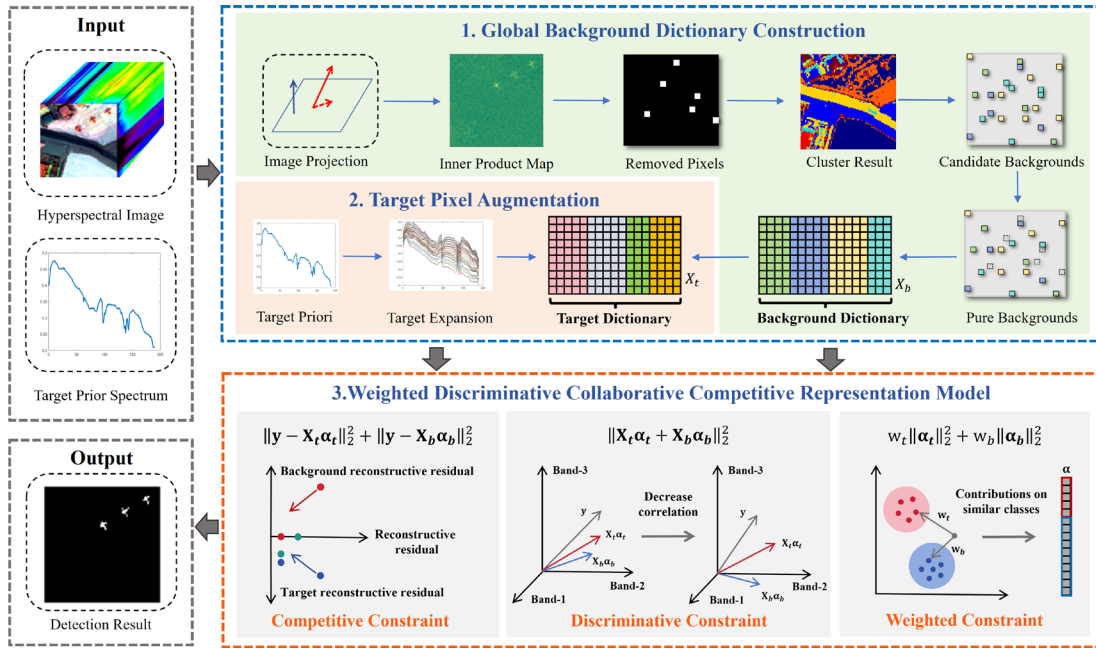


Fig. 1. Diagram of the proposed WDCCR for hyperspectral target detection. The model comprises three components: global background dictionary construction, target pixel augmentation, and the weighted discriminative collaborative competitive representation model. The first two components aim to construct a comprehensive dictionary for pixel representation and the third component utilizes an innovative representation model to represent the tested pixels using this dictionary.

the correlation between target and background reconstructions. The weight constraint assigns a similar class to the tested pixel more contributions in the representation model.

(2) We propose a pioneering approach to select pure background pixels for background dictionary construction, named the category-based pixel selection method with target orthogonal purification. Compared with existing methods, the proposed pixel selection method can avoid anomaly and noise contamination, and encompass a wide array of background feature types, contributing to a pure and complete background dictionary for the proposed representation learning model.

The remainder of this work is summarized as follows. Section II presents the proposed detector WDCCR, including global dictionary selection and weighted discriminative collaborative competitive representation. In Section III, experiments and analysis are performed on six datasets to verify the validity of the proposed detector. Finally, conclusions are provided in Section IV.

II. THE PROPOSED WDCCR METHOD

In this section, the proposed WDCCR method is described in detail. The WDCCR contains three main portions, i.e., background dictionary construction, target augmentation, and the weighted discriminative collaborative competitive representation model. Initially, background samples are selected from the entire image, forming the background dictionary by the category-based pixel selection method with target orthogonal purification (CPSTOP). Besides, target pixels are augmented using the selected background pixels and target priori. Subsequently, the weighted discriminative collaborative competitive representation model is employed to

compute the representation vector. Lastly, the detection result is determined based on the reconstructive residual difference between the background and the target. The diagram depicting the proposed WDCCR detector is presented in Fig. 1.

A. Global Background Dictionary Construction

Many existing representation learning methods rely on the dual concentric window strategy for constructing background dictionaries. However, their detection accuracy often fluctuates significantly with the window size. This poses challenges, particularly in densely populated target scenarios, as the fixed-size dual window approach fails to prevent genuine target pixels from being included. Consequently, this can lead to the generation of impure background dictionaries.

In attempts to address this issue, some methods focus on purifying the background dictionary. For instance, in [32], outlier pixels with values deviating more than double the standard deviation from the mean are eliminated from the background dictionary. Another strategy, as proposed in [33], considers pixels that cannot be adequately represented by the rest of the background and target spectra, assuming them to belong to the background. However, these approaches often remove abnormal pixels from the dual window without prior reference to targets, eliminating genuine background spectra potentially, which results in an insufficient representation of different land covers in backgrounds, ultimately deteriorating detection performance. Moreover, anomaly and noise pixel contamination is also a common failure of these methods.

To address these limitations, this study proposes a novel approach to select a global background dictionary, named category-based pixel selection with target orthogonal purification. The selection method aims to construct a pure

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

overcomplete background, avoiding the interference from both targets and noise, minimizing interference from both targets and noise while encompassing diverse ground objects representative of the background.

Consider the hyperspectral image $\mathbf{Y} \in \mathbb{R}^{L \times h \times w}$ with L bands, $h \times w$ is the spatial size. The n_t priori targets are given as $\mathbf{X}_t \in \mathbb{R}^{L \times n_t}$. The construction of the background dictionary can be described as follows:

- (1) Remove potential target pixels: first, potential target pixels are filtered out using target samples. The image \mathbf{Y} is projected into the orthogonal subspace of the target sample \mathbf{X}_t , followed by pixel-wise calculation of the inner product between the projected image and the original image. Pixels with smaller inner product values are closer to the sample, hence, the top five percent of pixels with the smallest inner product are removed.
- (2) Image cluster: to form a background dictionary covering elements from diverse land covers, we cluster the image \mathbf{Y} into ten classes through the K-means algorithm, where the number of pixels within each class is denoted as n_k .
- (3) Background pixel selection: for each class in the clustering result, after determining the class center, we compute the Euclidean distance between each pixel and its corresponding class center. Following this, while excluding pixels identified as potential targets, the nearest n_i pixels are selected from each class, as defined in Eq. (1), to construct the background dictionary.

$$n_i = \frac{n_k}{n_b} \quad (1)$$

where n_b is the predetermined total number of atoms.

Following the aforementioned procedures, the background dictionary \mathbf{X}_b consisting of n_b atoms is conducted under the guidance of priori. Compared to the popular dual window strategy, the construction method can well adapt to diverse hyperspectral scenes with different sizes and arrangements of targets. There is no need to set the window size for background construction and the analysis of the number of atoms will be given in the experiment section.

B. Target Pixel Augmentation

Due to the limited known target spectra for representation models, background samples are employed to augment target samples. In this paper, the target pixel augmentation method refers to the work in [49]. This augmentation involves incorporating a certain proportion θ of background spectra into the existing target spectrum, which can be formulated as follows.

$$\mathbf{s} = (1 - \theta) \times \mathbf{t} + \theta \times \mathbf{b}, 0 \leq \theta \leq 20\% \quad (2)$$

where \mathbf{t} is the reflectance of the prior target and \mathbf{b} is the reflectance of a selected background sample. We get the synthetic target spectrum vector by emulating the spectral variability with selected background samples. Finally, they are utilized to construct the target dictionary for the representation learning model in this work.

C. Weighted Discriminative Collaborative Competitive Representation Model

In HSIs target detection, collaborative representation is a widely used technique as it can alleviate the problem of spectral variability. Typical collaborative representation-based detectors are proposed based on such a technique [26], [32], [55]. Let $\mathbf{X} \in \mathbb{R}^{L \times N}$ denote the dictionary consisting of samples \mathbf{X}_t and \mathbf{X}_b , representing target and background samples, respectively. The objective is to solve a minimization problem by determining a weight vector α :

$$\min_{\alpha} \|\mathbf{y} - \mathbf{X}\alpha\|_2^2 \quad (3)$$

Incorporating a collaborative constraint term, the objective function can be reformulated as:

$$\argmin_{\alpha} \{\|\mathbf{y} - \mathbf{X}\alpha\|_2^2 + \lambda \|\alpha\|_2^2\} \quad (4)$$

where λ is a regularization parameter. As discussed in [51], the collaborative representation (CR) model emphasizes the contributions of different atoms, and it is the CR, rather than the l_1 -norm, that leads to improved classification accuracy.

However, existing CR models suffer from weak target background separability for several reasons. On the one hand, the CR model encourages all atoms to reconstruct the pixel. The optimization function aims to only minimize the overall residual. On the other hand, CR lacks promotion for utilizing the differential information from two classes. Besides, atoms from the target and background should not have an equivalent contribution during reconstruction. For HTD, a special binary classification problem, the above imperfections of CR perform more obvious ineffectiveness. The target pixels may be mixed with the reflectance of the surrounding background and the background itself contains a variety of elements so the reconstructions from target and background samples are close. Subsequently, the difference values of residuals can not be separated clearly. To bolster the separability between targets and background, this paper integrates three constraints into the collaborative representation model: the competitive constraint, discriminative constraint, and weighted constraint.

In HSI, when considering a tested pixel \mathbf{y} , the collaborative representation (CR) detector, as described in [26], aims to determine a weight vector α that minimizes the residual $r(\mathbf{y}) = \|\mathbf{y} - \mathbf{X}\alpha\|_2^2$. However, due to the subtle difference, the residual differences can not differentiate between the target and background effectively. To this end, a competitive constraint and discriminative constraint are imposed on the CR model.

The competitive constraint refers to the residual calculated separately for each class in addition to the residual computed for all samples and can be written as:

$$\|\mathbf{y} - \mathbf{X}_t \alpha_t\|_2^2 + \|\mathbf{y} - \mathbf{X}_b \alpha_b\|_2^2 \quad (5)$$

where α_t is a weight vector of the target, α_b is the weight vector of the background. The competitive constraint divides the residuals into target and background on the basis of the minimization of the entire residual of all samples. It can be reformulated as $\|\mathbf{y}\|_2^2 + \|\mathbf{X}_t \alpha_t\|_2^2 - 2\mathbf{y}^T(\mathbf{X}_t \alpha_t) + \|\mathbf{y}\|_2^2 + \|\mathbf{X}_b \alpha_b\|_2^2 - 2\mathbf{y}^T(\mathbf{X}_b \alpha_b)$. It's the minimization of residuals from individual classes that can promote the competition of

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

target and background samples. Define the correlation of two vectors \mathbf{x} and \mathbf{y} as $\text{Corr}(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$ since the dot product of \mathbf{x} and \mathbf{y} can reflect their geometrical and spatial correlation [52]. The correlation between the tested pixel \mathbf{y} and the representation with the target or background atoms can be expressed as $\text{Corr}(\mathbf{y}, \mathbf{X}_t \boldsymbol{\alpha}_t)$ and $\text{Corr}(\mathbf{y}, \mathbf{X}_b \boldsymbol{\alpha}_b)$. Through optimization, $\text{Corr}(\mathbf{y}, \mathbf{X}_t \boldsymbol{\alpha}_t)$ and $\text{Corr}(\mathbf{y}, \mathbf{X}_b \boldsymbol{\alpha}_b)$ are increased at the same time to achieve a competitive effect.

In contrast, the discriminative constraint designed for similar samples with variability in HSI can be expressed as:

$$\|\mathbf{X}_t \boldsymbol{\alpha}_t + \mathbf{X}_b \boldsymbol{\alpha}_b\|_2^2 \quad (6)$$

Equation (6) can be reformulated as $\|\mathbf{X}_t \boldsymbol{\alpha}_t\|_2^2 + \|\mathbf{X}_b \boldsymbol{\alpha}_b\|_2^2 + 2(\mathbf{X}_t \boldsymbol{\alpha}_t)^T (\mathbf{X}_b \boldsymbol{\alpha}_b)$, where the inner product $(\mathbf{X}_t \boldsymbol{\alpha}_t)^T (\mathbf{X}_b \boldsymbol{\alpha}_b)$ is defined as the correlation between representations $\mathbf{X}_t \boldsymbol{\alpha}_t$ and $\mathbf{X}_b \boldsymbol{\alpha}_b$, denoted by $\text{Corr}(\mathbf{X}_t \boldsymbol{\alpha}_t, \mathbf{X}_b \boldsymbol{\alpha}_b)$. By minimizing the discriminative term, the correlation is concurrently decreased and the reconstructions of target and background are pulled away, thereby enhancing discrimination between targets and backgrounds. In essence, if a tested pixel belongs to the target class, minimizing $\|\mathbf{X}_t \boldsymbol{\alpha}_t + \mathbf{X}_b \boldsymbol{\alpha}_b\|_2^2$ enhances the correlation $\text{Corr}(\mathbf{y}, \mathbf{X}_t \boldsymbol{\alpha}_t)$ and diminishes $\text{Corr}(\mathbf{y}, \mathbf{X}_b \boldsymbol{\alpha}_b)$. Consequently, the representation of target samples becomes more closely aligned with the tested pixel \mathbf{y} . This process strengthens inter-class discrimination, enabling each tested sample to obtain a discriminative and competitive representation for more effective detection.

The fourth part of the model is a weighted collaborative constraint. To fully leverage the discrimination between target and background samples, the varying importance of target and background for each query sample is incorporated in the representation. We assume that true class samples should contribute more to the representation, as a result, true class samples should carry smaller penalized weights during representation learning. Inspired by the k-means classification method, where a query sample belongs to the class with the smallest Euclidean distance, we calculate the average Euclidean distance between the tested pixel and samples from the target and background respectively to measure the probability of being the true class of the tested pixel. These average distances, denoted as w_b and w_t , are then utilized as weights for the representation coefficients of the background and target, respectively.

$$\begin{aligned} w_t \|\boldsymbol{\alpha}_t\|_2^2 + w_b \|\boldsymbol{\alpha}_b\|_2^2 \\ w_t = \frac{1}{n_b} \sum_{\mathbf{x} \in \mathbf{X}_t} \|\mathbf{y} - \mathbf{x}\|_2^2 \\ w_b = \frac{1}{n_b} \sum_{\mathbf{x} \in \mathbf{X}_b} \|\mathbf{y} - \mathbf{x}\|_2^2 \end{aligned} \quad (7)$$

According to the above three constraints, we have the final objective function:

$$\begin{aligned} \underset{\boldsymbol{\alpha}}{\text{argmin}} \{ \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 + \gamma(\|\mathbf{y} - \mathbf{X}_t \boldsymbol{\alpha}_t\|_2^2 + \|\mathbf{y} - \mathbf{X}_b \boldsymbol{\alpha}_b\|_2^2) \\ + \beta(\|\mathbf{X}_t \boldsymbol{\alpha}_t + \mathbf{X}_b \boldsymbol{\alpha}_b\|_2^2) + \lambda(w_b \|\boldsymbol{\alpha}_b\|_2^2 + w_t \|\boldsymbol{\alpha}_t\|_2^2) \} \end{aligned} \quad (8)$$

where $\|\cdot\|_2$ represents the l_2 norm. In the proposed WDCCR model, different classes participate in the representation. The separability between targets and backgrounds derived from residuals can be enhanced by integrating the three constraints, which helps effectively distinguish the subtle differences between the target and the background and contributes to good detection performance in hyperspectral target detection.

D. Analytic Solution for WDCCR Model

To compute the derivative with respect to $\boldsymbol{\alpha}$ in the objective function, Equation (8) can be reformulated as:

$$\underset{\boldsymbol{\alpha}}{\text{argmin}} \{ \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 + \gamma(\|\mathbf{y} - \bar{\mathbf{X}}_t \boldsymbol{\alpha}\|_2^2 + \|\mathbf{y} - \bar{\mathbf{X}}_b \boldsymbol{\alpha}\|_2^2) + \beta(\|\bar{\mathbf{X}}_t \boldsymbol{\alpha} + \bar{\mathbf{X}}_b \boldsymbol{\alpha}\|_2^2) + \lambda\|\mathbf{W}\boldsymbol{\alpha}\|_2^2 \} \quad (9)$$

where $\bar{\mathbf{X}}_t$ and $\bar{\mathbf{X}}_b$ are matrices of the same size as \mathbf{X} . $\bar{\mathbf{X}}_t$ and $\bar{\mathbf{X}}_b$ is defined as $[\mathbf{X}_t, 0, \dots, 0]$ and $[0, \dots, 0, \mathbf{X}_b]$, in which the location of $\bar{\mathbf{X}}_t$ and $\bar{\mathbf{X}}_b$ is the same as that in \mathbf{X} . According to the weight factor design, the weight matrix \mathbf{W} is defined as $\begin{bmatrix} \mathbf{W}_t & 0 \\ 0 & \mathbf{W}_b \end{bmatrix}$, with $\mathbf{W}_t = \text{diag}(\sqrt{w_t}, \sqrt{w_t}, \dots, \sqrt{w_t})$ and $\mathbf{W}_b = \text{diag}(\sqrt{w_b}, \sqrt{w_b}, \dots, \sqrt{w_b})$.

Let F_1 denotes the first part of the objective function $F_1 = \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2$. The second and third term with parameters γ and β is denoted as F_2 and F_3 . F_4 represents the l_2 -norm regularization $\lambda\|\mathbf{W}\boldsymbol{\alpha}\|_2^2$. The objective function is divided into four parts:

$$F_1 = \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 \quad (10)$$

$$F_2 = \gamma(\|\mathbf{y} - \bar{\mathbf{X}}_t \boldsymbol{\alpha}\|_2^2 + \|\mathbf{y} - \bar{\mathbf{X}}_b \boldsymbol{\alpha}\|_2^2) \quad (11)$$

$$F_3 = \beta(\|\bar{\mathbf{X}}_t \boldsymbol{\alpha} + \bar{\mathbf{X}}_b \boldsymbol{\alpha}\|_2^2) \quad (12)$$

$$F_4 = \lambda\|\mathbf{W}\boldsymbol{\alpha}\|_2^2 \quad (13)$$

Then each part of the model can be calculated by the following formulations.

$$\frac{dF_1}{d\boldsymbol{\alpha}} = \frac{d\|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2}{d\boldsymbol{\alpha}} = 2\mathbf{X}^T \mathbf{X} \boldsymbol{\alpha} - 2\mathbf{X}^T \mathbf{y} \quad (14)$$

$$\begin{aligned} \frac{dF_2}{d\boldsymbol{\alpha}} &= \gamma \frac{d(\|\mathbf{y} - \bar{\mathbf{X}}_t \boldsymbol{\alpha}\|_2^2 + \|\mathbf{y} - \bar{\mathbf{X}}_b \boldsymbol{\alpha}\|_2^2)}{d\boldsymbol{\alpha}} \\ &= 2\gamma [(\bar{\mathbf{X}}_t^T \bar{\mathbf{X}}_t + \bar{\mathbf{X}}_b^T \bar{\mathbf{X}}_b) \boldsymbol{\alpha} - (\bar{\mathbf{X}}_t^T + \bar{\mathbf{X}}_b^T) \mathbf{y}] \end{aligned} \quad (15)$$

$$\text{Define } \mathbf{M} \text{ as } \begin{bmatrix} \bar{\mathbf{X}}_t^T \bar{\mathbf{X}}_t & 0 \\ 0 & \bar{\mathbf{X}}_b^T \bar{\mathbf{X}}_b \end{bmatrix},$$

$$\frac{dF_2}{d\boldsymbol{\alpha}} = 2\gamma(\mathbf{M}\boldsymbol{\alpha} - \mathbf{X}^T \mathbf{y}) \quad (16)$$

$$\begin{aligned} \frac{dF_3}{d\boldsymbol{\alpha}} &= \beta \frac{d(\|\bar{\mathbf{X}}_t \boldsymbol{\alpha} + \bar{\mathbf{X}}_b \boldsymbol{\alpha}\|_2^2)}{d\boldsymbol{\alpha}} \\ &= 2\beta \mathbf{X}^T \mathbf{X} \boldsymbol{\alpha} \end{aligned} \quad (17)$$

$$\frac{dF_4}{d\boldsymbol{\alpha}} = \lambda \frac{d\|\mathbf{W}\boldsymbol{\alpha}\|_2^2}{d\boldsymbol{\alpha}} = 2\lambda \mathbf{W}^T \mathbf{W} \boldsymbol{\alpha} \quad (18)$$

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

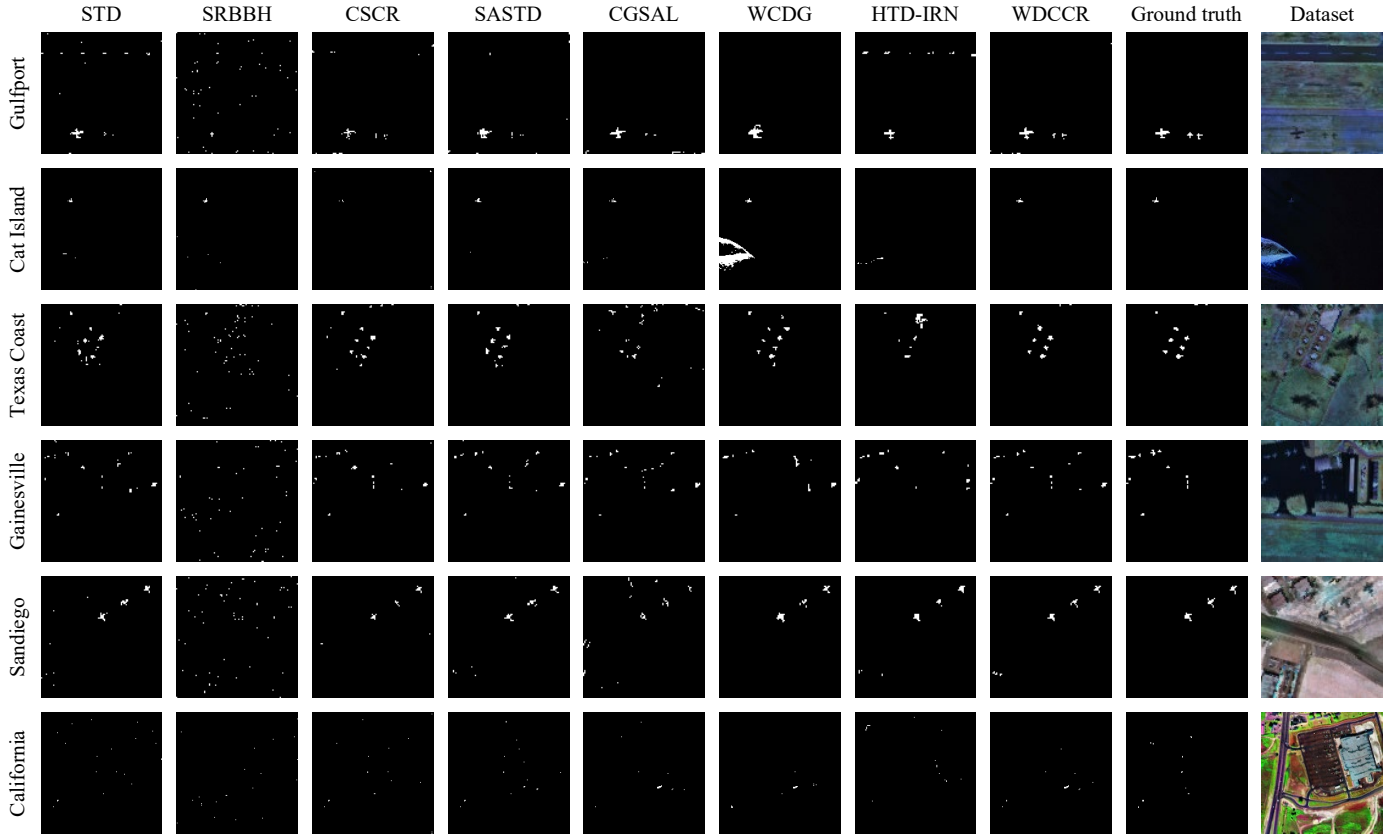


Fig. 2 The binary detection map comparison for six datasets.

Finally, consider the equation $\frac{dF_1}{d\alpha} + \frac{dF_2}{d\alpha} + \frac{dF_3}{d\alpha} + \frac{dF_4}{d\alpha} = 0$, where each term represents the derivative of a respective function with respect to α , Solving this equation analytically yields the solution for α :

$$\alpha = (1 + \gamma) \times ((1 + 2\beta)\mathbf{X}^T\mathbf{X} + \gamma\mathbf{M} + \lambda(\mathbf{W}^T\mathbf{W}))^{-1}\mathbf{X}^T\mathbf{y} \quad (19)$$

Since the representation vector α is obtained by a weighted discriminative collaborative competitive representation model, it can be decomposed into α_b and α_t components. The class-specific residuals r_t and r_b are calculated using target and background dictionary:

$$\begin{aligned} r_t &= \|\mathbf{y} - \mathbf{X}_t\alpha_t\|_2^2 \\ r_b &= \|\mathbf{y} - \mathbf{X}_b\alpha_b\|_2^2 \end{aligned} \quad (20)$$

Ultimately, the output of the proposed detector is expressed as:

$$D(\mathbf{y}) = r_b - r_t \quad (21)$$

The procedure of the proposed WDCCR algorithm is summarized in Algorithm I.

Algorithm I. WDCCR for Hyperspectral Target Detection

Input: hyperspectral data $\mathbf{Y} \in \mathbb{R}^{L \times h \times w}$, target samples $\mathbf{X}_t \in \mathbb{R}^{L \times n_t}$, parameters λ, β and γ .

Output: Two-dimensional plot of detection results.

Dictionary Construction and Augmentation:

1. Project the image \mathbf{Y} into the orthogonal subspace.
2. Calculate the inner product between the projected image and the original image.

3. Set the top five percent of pixels with the smallest inner product as the potential targets.
4. Cluster the image into ten classes by K-means.
5. Remove potential target pixels.
6. Select pixels with the nearest Euclidean distance to the class center from each class via Eq.(1).
7. Synthesize the target pixels via Eq. (2).

Optimized WDCCR Model:

8. Add three constraints in the CR model via Eq. (5)-(7).
9. Solve the optimization function Eq. (8).

Target Detection for HSI:

10. **For** each \mathbf{y} in image \mathbf{Y} ,
 - (1) calculate the weight vector via Eq. (19).
 - (2) calculate the residual r_b and r_t via Eq. (20).
 - (3) obtain detection statistics via (21).

End for

III. EXPERIMENTS AND ANALYSIS

In this section, extensive experiments conducted on six real-world datasets are used to validate the effectiveness of the proposed WDCCR detector. First of all, the hyperspectral datasets are described in detail. Then the target detection performance is compared against several state-of-the-art methods, including various representation-based and target detection algorithms. Subsequently, several parameter analyses are provided. Finally, the ablation study is conducted to ascertain the validity of the model.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

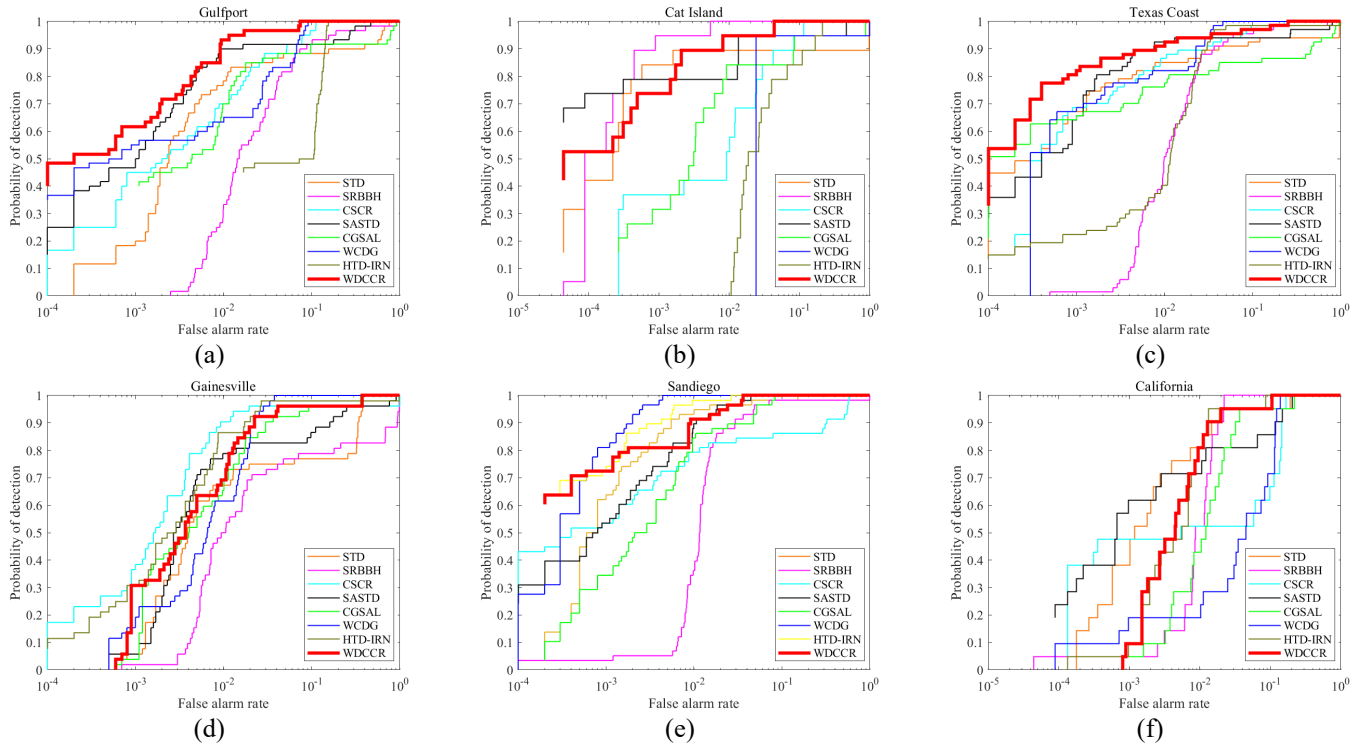


Fig. 3 ROC curves comparison of different detectors for three six hyperspectral datasets: (a) Gulfport; (b) Cat Island; (c) Texas Coast; (d) Gainesville; (e) San Diego; (f) California.

TABLE I. AUC VALUES COMPARISON OF DIFFERENT DETECTORS ON SIX DATASETS.

Method	STD	SRBBH	CSCR	SASTD	CGSAL	WCDG	HTD-IRN	WDCCR
Gulfport	0.9389	0.9577	0.9864	0.9772	0.9770	0.9813	0.9365	0.9987
Cat Island	0.9231	0.9997	0.9834	0.9857	0.9660	0.9611	0.9606	0.9993
Texas Coast	0.9440	0.9776	0.9915	0.9738	0.9604	0.9940	0.9802	0.9903
Gainesville	0.9094	0.8412	0.9777	0.9523	0.9668	0.9892	0.9855	0.9919
San Diego	0.9963	0.9776	0.9380	0.9963	0.9693	0.9994	0.9989	0.9977
California	0.9921	0.9900	0.9459	0.9779	0.9715	0.9456	0.9928	0.9950

A. Dataset Description

The datasets used in this work contain various scenes, including airports, beaches, urban areas, and diverse target objects. The first dataset pertains to Gulfport, USA, featuring HSI measuring 100×100 pixels with 191 bands. The sensor's spatial resolution is $7.1m$, covering spectral bands ranging from $0.4\text{--}2.5\mu m$. The second dataset was acquired from Cat Island, Japan, with a spatial size of 150×150 pixels, and 188 spectral bands. Its spatial resolution is $17.2m$. Both the third and fourth datasets capture urban scenes. The Texas Coast dataset has 204 spectral bands, and the Gainesville dataset has 191 spectral bands, with corresponding spatial resolutions of $17.2m$ and $3.5m$, respectively. They both cover wavelengths from 0.4 to $2.5\mu m$. The fifth dataset, acquired from San Diego, CA, USA, by AVIRIS, features a spatial resolution of $3.5m$. It consists of 224 spectral bands covering wavelengths from

$0.37\text{--}2.5\mu m$. After removing poor-quality bands, 189 bands were retained for experiments. The sixth dataset was collected by the Hyperspectral Digital Image Collection Experiment (HYDICE) sensor in California. It comprises 150×150 pixels with 210 spectral bands. The spatial resolution is $2m$, and the spectral resolution is $0.01\mu m$. 162 bands were retained for experiments. Fig. 2. displays the pseudo-color images and ground-truth maps of all datasets.

B. Experimental Settings

(1) Performance metrics

To assess the performance of the WDCCR method for target detection, three widely recognized evaluation criteria are adopted: receiver operating characteristic (ROC) curve, area under the ROC curve (AUC) value, and target-background separability (TBS) map.

The ROC curve provides insight into performance across

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

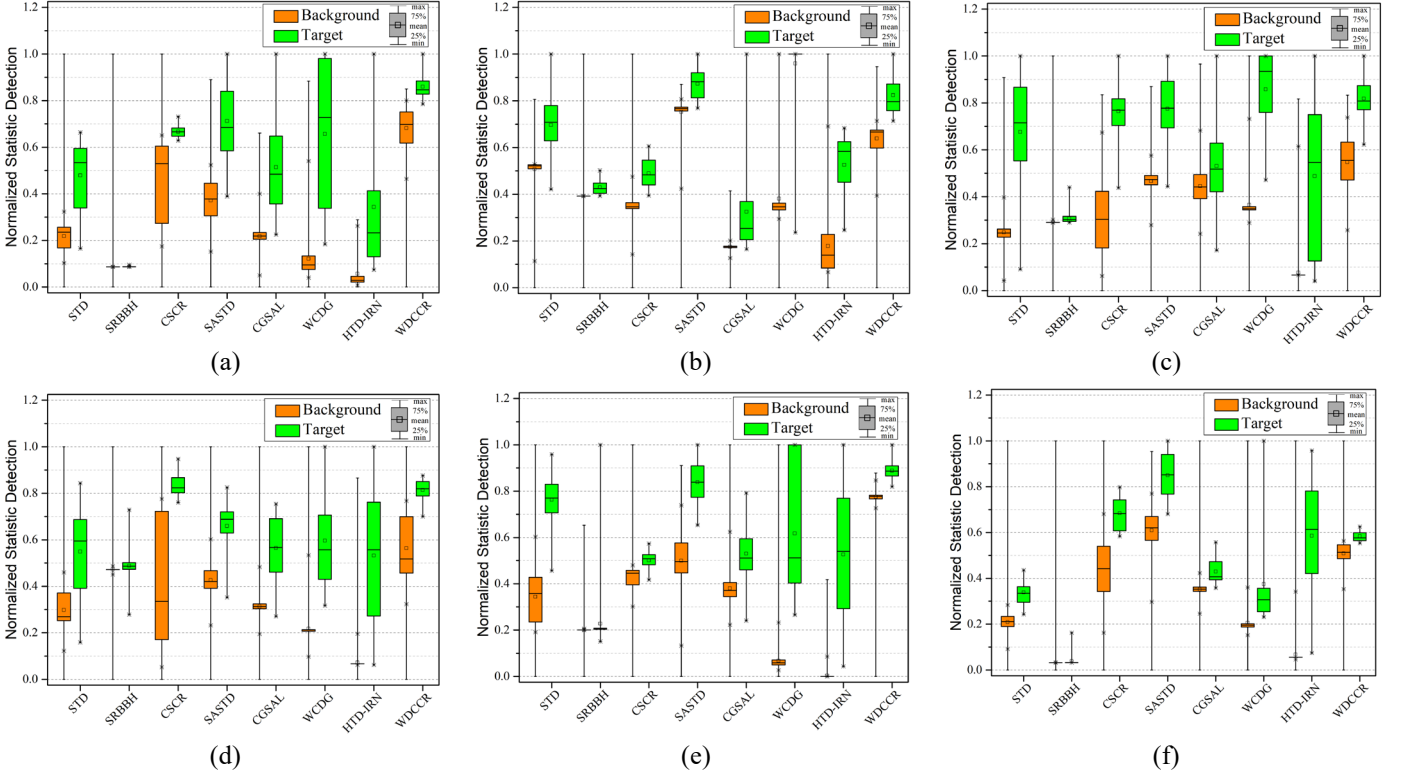


Fig. 4. Target background separability map comparison of different detectors for six hyperspectral datasets. (a) Gulfport. (b) Cat Island. (c) Texas Coast, (d) Gainesville, (e) San Diego, (f) California.

various thresholds by plotting the probability of detection (PD) against the false alarm rate (FAR) under fixed conditions [29]. PD is calculated by the ratio of the detected number of detected target pixels to that of real target pixels, while the ordinate FAR refers to the false alarm rate, calculated by the ratio of the number of misjudged pixels to all pixels in the image. Generally, a ROC curve's proximity to the top-left corner of the plot signifies better performance.

AUC, derived from the ROC curve, offers a quantitative evaluation. It is the area under the ROC curve ranging between 0 and 1 [58]. A higher AUC value corresponds to superior detector performance.

The TBS map evaluates the separability between targets and backgrounds through box plots. The ordinate is expressed by the normalized detection statistic [59-61]. In this map, pixels with detection results are divided into target and background classes based on their labels. The normalized detection statistic is then calculated for each class and visualized as a box plot. A box plot illustrates the main body of the normalized detection statistic ranging from 25% to 75% [29]. A larger interval between the two box plots typically indicates better performance, highlighting targets while minimizing background interference.

(2). Comparison detectors and parameter settings

Seven state-of-the-art hyperspectral target detectors are employed to compare the detection performance of the method. The compared methods include five detectors, i.e., STD [24], SRBBH [25], CSCR [26], SASTD [27], WCDG [28], which

are based on representation learning, a detector based on spectral linear mixing model CGSAL [29], and the HTD-IRN [53] that is constructed based on deep learning. To ensure fairness, the same a priori spectrum was provided for all methods.

It should be noted that for methods based on representation learning, including STD, SRBBH, CSCR, and SASTD, the parameter settings for dual window sizes are crucial and tailored to the target's shape. The selection of these window sizes significantly influences the detector's performance. Optimal window sizes are investigated through the SRBBH method, and these were subsequently applied to other methods requiring similar window size adjustments to ensure optimal results. For the CGSAL method, in addition to the a priori spectrum, background endmembers are required. In this paper, we cluster the HSI and calculate the center of each class. The pixel with the closest Euclidean distance to the class center is chosen as the background endmember. The window size in WCDG is the same as the inner window of SRBBH. In HTD-IRN, both endmember number and the regularization parameter of sparse loss are set to 10 for datasets except for San Diego, whose endmember number is set to 30. The size of [inner window, outer window] for six datasets are set as [13,17], [13,21], [9,11], [13,15], [13,15], and [7,11] respectively.

For the parameters of the proposed WDCCR, the number of cluster classes is set to 10. The parameters λ , β , and γ are set

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

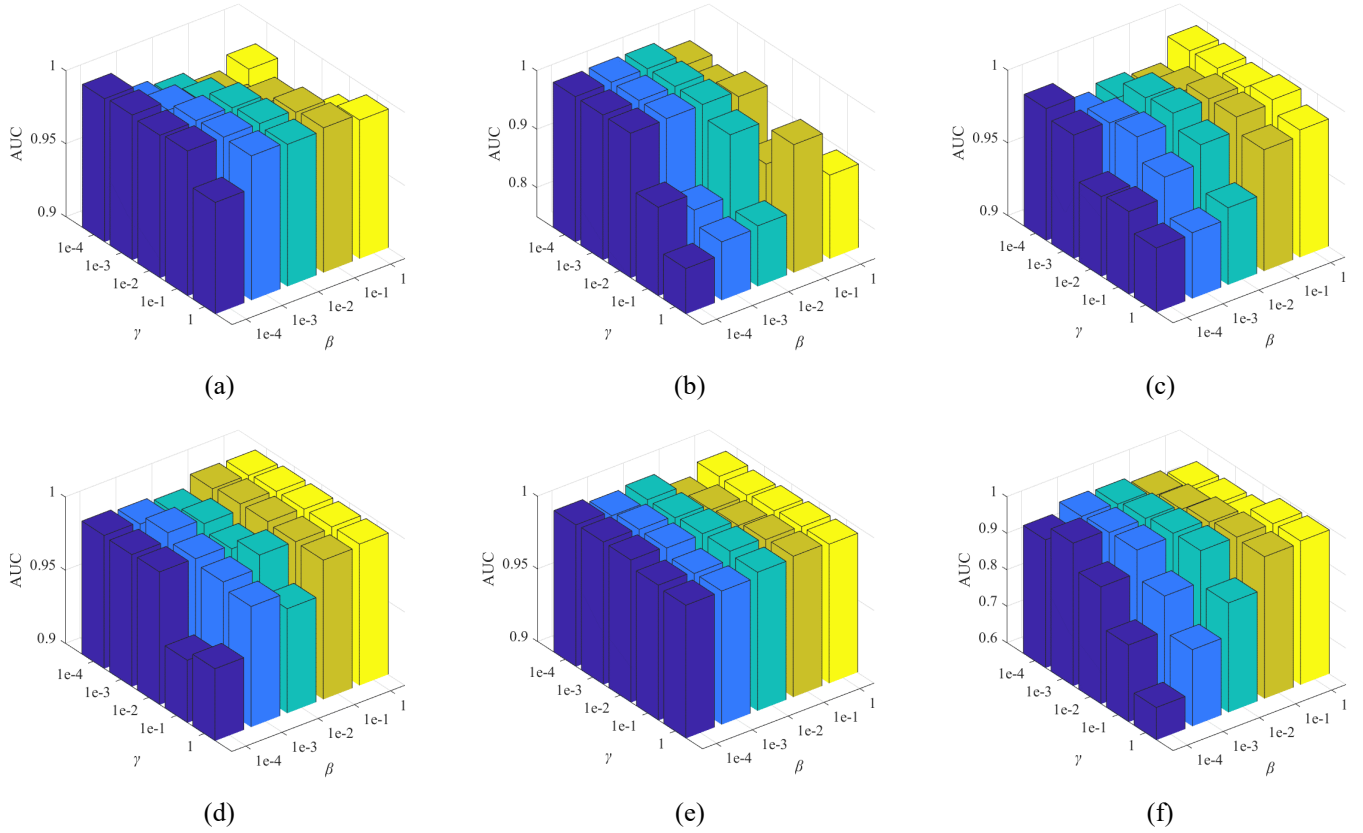


Fig. 5. AUC value performance with different coefficients β and γ . (a) Gulfport. (b) Cat Island. (c) Texas Coast, (d) Gainesville, (e) San Diego, (f) California.

to $1e-2$. Additionally, the number of atoms was set to 200. All experiments were conducted on a 13th Gen Intel(R) Core (TM) i5-13400 CPU machine running a 64-bit Windows 11 Operating System (OS), using MATLAB 2023a software.

C. Detection performance

The binarized detection maps are displayed in Fig. 2. Since the detector outputs maps with statistical detection, a thresholding process is necessary to convert them into a binary format [29]. In the case of having n_r target pixels in the image, we sort the detection results in descending order, and then define the value of the n_r th pixel in descending order as the threshold. Pixels with values larger than the threshold are assigned a value of 1, while those smaller than that are assigned a value of 0. An ideal detector is characterized not only by a high AUC value but also by producing a binary map that closely aligns with the ground truth map, with minimal false alarm pixels. In the Gulfport dataset, three airplanes are located at the bottom of the image. Methods such as STD and HID-IRN mistakenly detect pixels in spaced rows at the top of the image. CSCR, SASTD, and CGSAL exhibit bits of false alarm pixels and fail to completely detect the target. WCDG, although devoid of false alarm pixels outside the target location, detects a target larger than the actual one. In the Texas Coast dataset, CGSAL, WCDG, and HTD-IRN struggle to detect targets lie in the left column with smaller sizes. Moving to the Cat Island dataset, WCDG mistakenly detects

submerged rocks in the image, unrelated to our target of interest. SRBBH struggles to accurately locate the target across different datasets, especially having irregular, small but widely distributed false alarms. Similarly, the STD, CSCR, SASTD, and CGSAL methods encounter challenges akin to those observed in the Gulfport dataset. Conversely, WDCCR's target distribution closely aligns with the real target corresponding to ground truth. In the California dataset, the vehicle targets have the smallest size among the six datasets, they are hard to locate in binary maps relatively. Concluding the binary maps above, our proposed WDCCR can detect intact targets with as few false alarms as possible. Since we discard the dual window strategy and adapt the CPSTOP for dictionary construction, the misdetection around the target can be avoided. Besides, background can be suppressed successfully due to the competitive constraint. It minimizes the decisional residuals of the competitive class so that the background has small reconstruction residuals correctly.

In the comparison of ROC curves, as illustrated in Fig. 3, across the Gulfport, Texas Coast, and San Diego datasets, detectors such as SRBBH, CSCR, CGSAL, and HTD-IRN struggle to achieve a high PD at the low FAR when compared to the proposed WDCCR method. Especially for the Gulfport dataset, WDCCR consistently maintains a leading position across varying false alarm rates. Some of these detectors employ a fixed dual window, which may yield suboptimal results when the target objects are densely arranged or exhibit

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

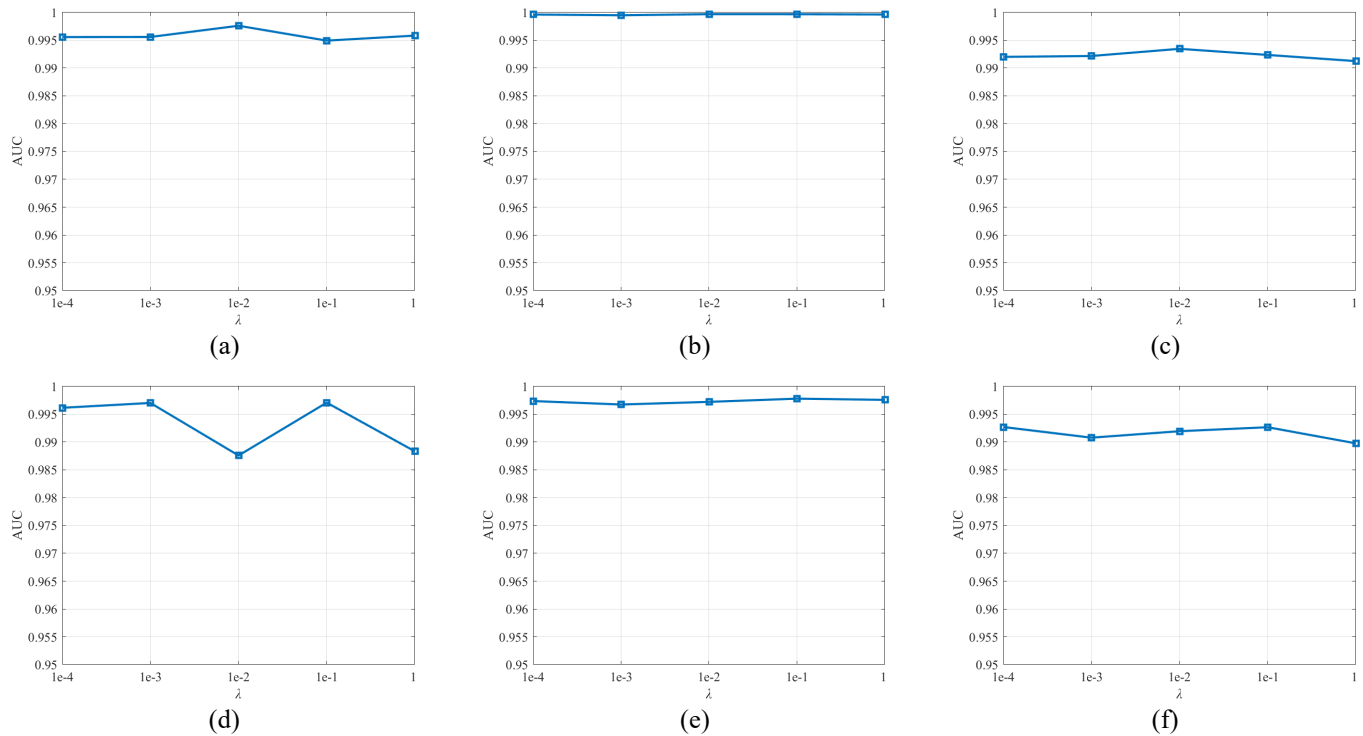


Fig. 6. AUC value with different λ . (a) Gulfport. (b) Cat Island. (c) Texas Coast, (d) Gainesville, (e) San Diego, (f) California.

significant shape variations. In the San Diego dataset, although WDCCR does not outperform STD, WCDG, and HTD-IRN with FAR lower than 0.01, it demonstrates superior detection performance, reaching a detection probability of 1 when the false alarm rate exceeds 0.05, outperforming other detectors. Owing to the weight constraints giving different classes larger penalization in representation, the proposed method detects more target pixels under the same false alarm rate. Originally vague pixels get more contribution from priori target and have smaller target residuals. Overall, WDCCR generally exhibits superior performance at low FAM compared to other detectors, with its ROC curve positioned closer to the top left corner, indicating its effectiveness in target detection for HSIs.

The detector performance is further quantitatively assessed using AUC. Table I showcases the AUC values of different detectors. The bold number with underline means the best AUC result among different methods and the bold number represents the second highest AUC. The AUC values for WDCCR across all datasets exceed 0.99, indicating superior performance, particularly evident in the Gulfport and Gainesville datasets. While our detection results may not consistently rank as the highest in other datasets, they still surpass the 0.99 accuracy. Notably, the proposed WDCCR consistently delivers stable detection results across different datasets due to its alternating dictionary selection and detection model, setting it apart from other state-of-the-art methods. The TBS maps are presented in Fig. 4., where the orange and green boxes represent background and target detection statistics, respectively. In the Gulfport and Gainesville datasets, SRBBH shows a limited detection range, which compromises its separability capability, particularly

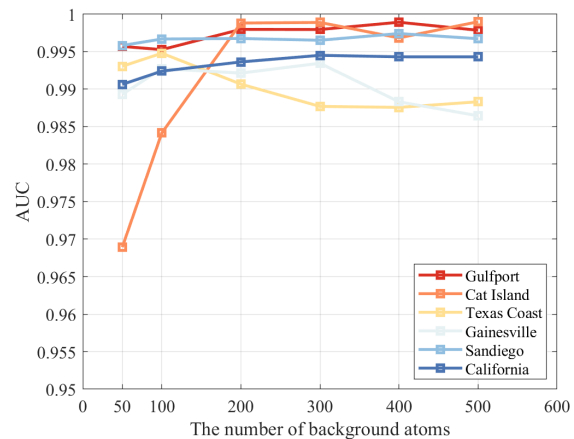


Fig. 7. AUC comparison of different numbers of background atoms in WDCCR for six datasets.

noticeable in the Gulfport and Gainesville datasets. In the Cat Island dataset, the majority method can limit the normalized statistic detection to a small range for the reason that the background has simple components. SASTD effectively highlights target statistics but struggles to suppress the background range adequately, resulting in suboptimal performance in complex scenes such as Gainesville and California. STD and HTD-IRN yield similar separability results with well-suppressed background boxes; however, they exhibit overlap with the 0%-25% target value range, potentially leading to false alarm detections. CGSAL displays significant overlap on the Texas Coast dataset, possibly influenced by the endmember extraction process. In contrast,

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

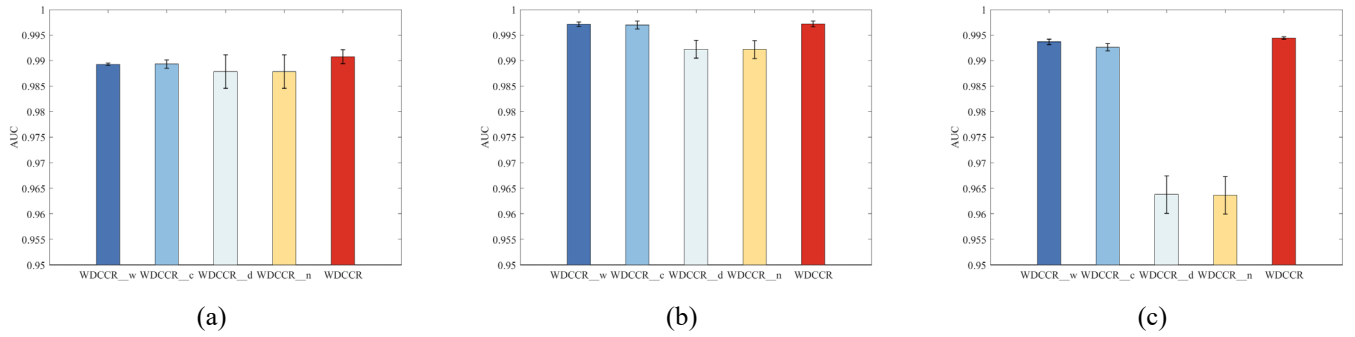


Fig. 8. The AUC comparison of different ablation experiments on three datasets: (a) Gainesville, (b) Sandiego, and (c) California.

the WDCCR demonstrates satisfactory TBS performance by indicating an obvious gap between the target and background ranges. WDCCR achieves this by constructing a more comprehensive background dictionary for various land cover types and using a more discriminative model. The discriminative constraint plays a part in decreasing the correlation between the target and background class during reconstructions. It pulls away the residuals of two classes, thus enhancing its target-background separability. Besides, the CPSTOP strategy to construct a dictionary can suppress the background with the target highlighted.

D. Analysis of parameters

(1) Analysis of β and γ

This section delves into the influence of coefficients β and γ , corresponding to the competitive and discriminative constraint, respectively. Given that these two constraints hold a similar level of importance, apart from the weight constraint, we initially fix λ and conduct experiments by setting the values of both β and γ as 1e-4, 1e-3, 1e-2, 1e-1, 1. The results are shown in Fig. 5. The AUC results are affected greatly by β on several datasets but do not exhibit a consistent pattern of change with variations in parameters β across different datasets, and the same is true for parameter γ .

In urban scenes across three datasets, e.g., Texas Coast, Gainesville, and California, the AUC value fails to reach an ideal level when β is too small. This phenomenon arises due to several objects in urban environments, such as roofs and roads, which may lack discriminative features, necessitating the role of the discriminative term associated with β in the model. Conversely, the Cat Island dataset yields low AUC values when either β or γ is large, attributable to the imbalanced distribution of land covers, which occupies a majority of the image area, with ship target pixels constituting only a minor portion. The reason is that too large a weight on the extra competitive and discriminative term will impede the aim of representation learning of minimizing the residuals.

In contrast, parameters β and γ have a small influence on Gulfport and Sandiego due to significant differences in land covers. The AUC values under varying β and γ fluctuate above 0.95. Notably, the AUC remains consistently high when $\beta = 0.01$ and $\gamma = 0.01$. Therefore, it is recommended to set $\beta = 0.01$ and $\gamma = 0.01$ for the WDCCR detector to get a

robust solution across different datasets and achieve optimal detection performance.

(2) Analysis of λ

This section investigates the sensitivity of the parameter λ . We set λ values as 1e-4, 1e-3, 1e-2, 1e-1, and 1, while maintaining $\beta = 0.01$ and $\gamma = 0.01$. The AUC results under the influence of λ are displayed in Fig. 6. Compared to the parameters β and γ , the proposed model is insensitive to λ , and the AUC performance tends to be stable when λ is no larger than 0.01. When λ is larger than 0.01, the AUC has a slight downward trend. In the Gainesville and California datasets, the AUC fell below 0.9 at $\lambda = 1$. Since λ represents the parameter of the regularization constraint rather than the reconstruction residuals, it is assigned a small value to maintain the interpretability of the function. Overall, WDCCR has robustness with the variation of λ . It has a slight influence on detection results which the AUC undulates within 0.01. Consequently, λ is recommended to be set as 1e-2 for different datasets to achieve optimal performance.

(3) Analysis of the number of atoms

In this section, we delve into the influence of the background atom number. The number of atoms is set to 50, 100, 200, 300, 400, and 500 to compare the detection results across different atom numbers.

In the Cat Island dataset, there is a noticeable decline in AUC value when the background atoms are less than 200. This decline can be attributed to the predominant presence of sea and bank areas with reefs in the image scene. Insufficient atoms fail to capture the variability of the land covers, which constitutes a small portion of the image. Conversely, in datasets such as Texas Coast and Gainesville, characterized by urban scenes, an excess of atoms takes the risk of containing noise pixels, such as cloud cover and shadows. As shown in the AUC plot, the AUC decreases below 0.99 when the number of atoms exceeds 200 and 300. For military targets in Gulfport and Sandiego, the number of atoms has a tiny influence on AUC due to the widespread distribution of background land covers.

Summarizing the results in Fig. 7, when the atom number is smaller than 200, the AUC value fails to reach an excellent effect across all datasets. Conversely, when the atom number

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

exceeds 200, the AUC value shows no obvious increase and may even decrease for Gainesville. Consequently, the optimal and necessary number of atoms is determined to be 200.

(4) Analysis of three constraints

To validate the effectiveness of the competitive constraint, discriminative constraint, and weight factor in the proposed WDCCR detector, we conduct ablation experiments by removing each of these components individually and comparing the detection results with the unabridged model. Abbreviations are used to represent different ablation experiments. "WDCCR_c", "WDCCR_d", and "WDCCR_w" stand for experiments without the competitive term, discriminative term, and weight term, respectively, "WDCCR_n" stands for experiments without both the competitive and discriminative terms, while "WDCCR" represents the full model. The comparative results on Gainesville, Sandiego, and California datasets are displayed in Fig. 8. It can be seen that the full model can achieve robust detection performance and outperform its competitors, which confirms the effectiveness of all three constraints. The competitive constraint and the discriminative constraint play a more important role in the model. The constraint decreases the correlation of target and background in the progress of reconstruction, effectively discriminating them with less similarity.

V. CONCLUSION

In this work, we propose a novel target detector for hyperspectral images, named the weighted discriminative collaborative competitive representation model (WDCCR) with global background dictionary construction. The model employs a new category-based pixel selection method with target orthogonal purification to form a pure and complete background dictionary, which can avoid anomaly and noise contamination. To further distinguish between the target and background, the discriminative constraint, the competitive constraint, and the weighted constraint are integrated into the model. Compared with deep-learning methods, the background dictionary construction method and different constraints in the representation model are designed to solve the problem of poor separability between target and background without training. A series of experiments are conducted on six real-world datasets to confirm the effectiveness of the model. These experimental results including three common evaluation criteria and binary maps show that the WDCCR detector has superb and robust performance compared to other advanced target detection methods. Future works will be performed to further enhance the performance of subpixel target detection based on discriminative deep learning-based methods.

REFERENCES

- [1] R. O. Green et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)", *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227-248, Sep. 1998.
- [2] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications", *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29-43, 2002.
- [3] D. Landgrebe, "Hyperspectral image data analysis", *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17-28, Jan. 2002.
- [4] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges", *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6-36, Jun. 2013.
- [5] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation", *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12-16, Jan. 2002.
- [6] S. Mei, Z. Han, M. Ma, F. Xu, and X. Li, "A Novel Center-Boundary Metric Loss to Learn Discriminative Features for Hyperspectral Image Classification", *IEEE Trans. Geosci. Remote Sens.*, vol. 62, pp. 1-16, 2024.
- [7] J. Liu, Z. Hou, W. Li, R. Tao, D. Orlando, and H. Li, "Multipixel anomaly detection with unknown patterns for hyperspectral imagery", *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5557-5567, October 2022.
- [8] X. Hu et al., "Spatial-spectral identification of abnormal leukocytes based on microscopic hyperspectral imaging technology", *J. Innov. Opt. Health Sci.*, vol. 13, no. 2, Mar. 2020.
- [9] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Detection algorithms in hyperspectral imaging systems", *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 24-33, Jan. 2014.
- [10] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications", *J. Lincoln Lab.*, vol. 14, no. 1, pp. 79-116, 2003.
- [11] N. M. Nasrabadi, "Hyperspectral target detection: An overview of current and future challenges", *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 34-44, Jan. 2014.
- [12] C.-I. Chang, "An information-theoretic approach to spectral variability similarity and discrimination for hyperspectral image analysis", *IEEE Trans. Inf. Theory*, vol. 46, no. 5, pp. 1927-1932, Aug. 2000.
- [13] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic", *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427-438, Feb. 2005.
- [14] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Is there a best hyperspectral detection algorithm?", *Proc. SPIE*, vol. 7334, Apr. 2009.
- [15] W. Farrand, "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.*, vol. 59, no. 1, pp. 64-76, Jan. 1997.
- [16] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208-216, Jan. 1992.
- [17] Chein-I Chang, "Orthogonal subspace projection (OSP) revisited: a comprehensive study and analysis", *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 502-518, Mar. 2005.
- [18] J. Peng, W. Sun and Q. Du, "Self-paced joint sparse representation for the classification of hyperspectral images", *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1183-1194, Feb. 2019.
- [19] J. Peng, L. Li, and Y. Y. Tang, "Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images", *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 6, pp. 1790-1802, Jun. 2019.
- [20] X. Jiao and C.-I. Chang, "Kernel-based constrained energy minimization (K-CEM)", *Proc. SPIE*, vol. 6966, pp. 523-533, Jul. 2008.
- [21] H. Kwon and N. M. Nasrabadi, "Kernel spectral matched filter for hyperspectral imagery", *Int. J. Comput. Vis.*, vol. 71, no. 2, pp. 127-141, Feb. 2007.
- [22] L. Capobianco, A. Garzelli and G. Camps-Valls, "Target detection with semisupervised kernel orthogonal subspace projection", *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3822-3833, Nov. 2009.
- [23] L. Zhang and L. Zhang, "Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities", *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 270-294, Jun. 2022.
- [24] Y. Chen, N. M. Nasrabadi and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery", *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 629-640, Jun. 2011.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

- [25] Y. Zhang, B. Du and L. Zhang, "A sparse representation-based binary hypothesis model for target detection in hyperspectral images", *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1346-1354, Mar. 2015.
- [26] W. Li, Q. Du, and B. Zhang, "Combined sparse and collaborative representation for hyperspectral target detection", *Pattern Recognit.*, vol. 48, no. 12, pp. 3904-3916, Dec. 2015.
- [27] Y. Zhang, B. Du, Y. Zhang, and L. Zhang, "Spatially adaptive sparse representation for target detection in hyperspectral images", *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 1923-1927, Nov. 2017.
- [28] X. Zhao, W. Li, C. Zhao, and R. Tao, "Hyperspectral target detection based on weighted Cauchy distance graph and local adaptive collaborative representation", *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
- [29] D. Zhu, B. Du, M. Hu, Y. Dong, and L. Zhang, "Collaborative-guided spectral abundance learning with bilinear mixing model for hyperspectral subpixel target detection", *Neural Netw.*, vol. 163, pp. 205-218, Jun. 2023.
- [30] X. Lu, W. Zhang, and X. Li, "A hybrid sparsity and distance-based discrimination detector for hyperspectral images", *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1704-1717, Mar. 2018.
- [31] C. Li, W. Zhang, Y. Zhang, Z. Chen and H. Gao, "Adaptively dictionary construction for hyperspectral target detection", *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1-5, 2023.
- [32] D. Zhu, B. Du and L. Zhang, "Binary-class collaborative representation for target detection in hyperspectral images", *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 7, pp. 1100-1104, Jul. 2019.
- [33] Q. Ling, K. Li, Z. Li, Z. Lin, and J. Wang, "Hyperspectral detection and unmixing of subpixel target using iterative constrained sparse representation", *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 1049-1063, 2022.
- [34] T. Guo, F. Luo, L. Zhang, B. Zhang, X. Tan, and X. Zhou, "Learning structurally incoherent background and target dictionaries for hyperspectral target detection", *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3521-3533, 2020.
- [35] L. Zhang, L. Zhang, D. Tao, X. Huang and B. Du, "Hyperspectral remote sensing image subpixel target detection based on supervised metric learning", *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4955-4965, Aug. 2014.
- [36] Y. Dong, B. Du and L. Zhang, "Target detection based on random forest metric learning", *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 4, pp. 1830-1838, Apr. 2015.
- [37] S. Meerdink et al., "Multitarget multiple-instance learning for hyperspectral target detection", *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
- [38] C. Jiao, B. Yang, Q. Wang, G. Wang, and J. Wu, "Discriminative multiple-instance hyperspectral subpixel target characterization", *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022.
- [39] L. Zhang, L. Song, B. Du and Y. Zhang, "Nonlocal low-rank tensor completion for visual data", *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 673-685, Feb. 2021.
- [40] X. Zhao, K. Liu, K. Gao, and W. Li, "Hyperspectral time-series target detection based on spectral perception and spatial-temporal tensor decomposition", *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023.
- [41] K. Sun, X. Geng and L. Ji, "A new sparsity-based band selection method for target detection of hyperspectral image", *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 329-333, Feb. 2015.
- [42] Y. Xu, Q. Du, and N. H. Younan, "Particle swarm optimization-based band selection for hyperspectral target detection", *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 554-558, Apr. 2017.
- [43] X. Liu, S. Bourennane and C. Fossati, "Reduction of signal-dependent noise from hyperspectral images for target detection", *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5396-5411, Sep. 2014.
- [44] F. Li and M. Song, "Sequential band fusion for hyperspectral target detection", *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1-24, 2022.
- [45] Y. Cai, X. Liu, and Z. Cai, "BS-Nets: An end-to-end framework for band selection of hyperspectral image", *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 1969-1984, Mar. 2020.
- [46] Q. Wang, Z. Yuan, Q. Du, and X. Li, "GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection", *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3-13, Jan. 2019.
- [47] J. Du and Z. Li, "A hyperspectral target detection framework with subtraction pixel pair features", *IEEE Access*, vol. 6, pp. 45562-45577, 2018.
- [48] G. Zhang, S. Zhao, W. Li, Q. Du, Q. Ran, and R. Tao, "HTD-Net: A deep convolutional neural network for target detection in hyperspectral imagery", *Remote Sens.*, vol. 12, no. 9, pp. 1489, 2020.
- [49] D. Zhu, B. Du and L. Zhang, "Two-stream convolutional networks for hyperspectral target detection", *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6907-6921, Aug. 2021.
- [50] W. Xie, J. Lei, J. Yang, Y. Li, Q. Du, and Z. Li, "Deep Latent Spectral Representation Learning-Based Hyperspectral Band Selection for Target Detection", *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 2015-2026, 2020.
- [51] H. Qin, W. Xie, Y. Li, and Q. Du, "HTD-VIT: Spectral-spatial joint hyperspectral target detection with vision transformer", *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, pp. 1967-1970, Jul. 2022.
- [52] X. Yang, M. Zhao, S. Shi, and J. Chen, "Deep constrained energy minimization for hyperspectral target detection", *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 8049-8063, 2022.
- [53] D. Shen, X. Ma, W. Kong, J. Liu, J. Wang, and H. Wang, "Hyperspectral target detection based on interpretable representation network", *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1-16, 2023.
- [54] J. Zhao, G. Wang, B. Zhou, J. Ying, and J. Liu, "SRA-CEM: An Improved CEM Target Detection Algorithm for Hyperspectral Images Based on Subregion Analysis," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 6026-6037, Jun. 2023.
- [55] W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection", *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1463-1474, Mar. 2015.
- [56] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?", *Proc. Int. Conf. Comput. Vis.*, pp. 471-478, Nov. 2011.
- [57] J. Gou, L. Wang, Z. Yi, Y. Yuan, W. Ou, and Q. Mao, "Weighted discriminative collaborative competitive representation for robust image classification", *Neural Netw.*, vol. 125, pp. 104-120, May 2020.
- [58] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions", *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251-255, Apr. 2008.
- [59] S. Khazai, S. Homayouni, A. Safari, and B. Mojaradi, "Anomaly detection in hyperspectral images based on an adaptive support vector method", *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 646-650, Jul. 2011.
- [60] Z. Hou, W. Li, L. Li, R. Tao, and Q. Du, "Hyperspectral change detection based on multiple morphological profiles", *IEEE Trans. Geosci. Remote Sens.*, Jul. 2021.
- [61] D. Zhu, B. Du, Y. Dong, and L. Zhang, "Target Detection with Spatial-Spectral Adaptive Sample Generation and Deep Metric Learning for Hyperspectral Imagery," *IEEE Trans. Multimedia*, vol. 25, pp. 6538-6550, 2023.