Investigating the Effects of a Combined Spatial and Spectral Dimensionality Reduction Approach for Aerial Hyperspectral Target Detection Applications

Fraser Macfarlane [1]*[ID]**, Paul Murray** [1]**, Stephen Marshall** [1] **and Henry White** [2]

[1]   Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, United Kingdom; {fraser.macfarlane, paul.murray, stephen.marshall}@strath.ac.uk

[2]   BAE Systems, Air Sector, Filton, Bristol, UK

*   Correspondence: fraser.macfarlane@strath.ac.uk

Abstract: Target detection and classification is an important application of hyperspectral imaging in remote sensing. A wide range of algorithms for target detection in hyperspectral images have been developed in the last few decades. Given the nature of hyperspectral images, they exhibit large quantities of redundant information and are therefore compressible. Dimensionality reduction is an effective means of both compressing and denoising data. Although spectral dimensionality reduction is prevalent in hyperspectral target detection applications, the spatial redundancy of a scene is rarely exploited. By applying simple spatial masking techniques as a preprocessing step to disregard pixels of definite disinterest, the subsequent spectral dimensionality reduction process is simpler, less costly and more informative. This paper proposes a processing pipeline to compress hyperspectral images both spatially and spectrally before applying target detection algorithms to the resultant scene. The combination of several different spectral dimensionality reduction methods and target detection algorithms, within the proposed pipeline, are evaluated. We find that the Adaptive Cosine Estimator produces an improved F1 score and Mathews Correlation Coefficient when compared to unprocessed data. We also show that by using the proposed pipeline the data can be compressed by over 90% and target detection performance is maintained.

1. Introduction

Remote sensing from aerial and satellite platforms has become increasingly prevalent and is an important source of information in areas of research including disaster relief [1], determining land usage [2] and assessing vegetation health [3]. Remote sensing platforms are also often deployed in military and security applications such as change detection [4,5], target tracking [6] and classification. Target Detection (TD) from airborne imagery is a major challenge and active area of research within the disciplines of signal and image processing [7–9]. There have been a wide range of TD algorithms of varying complexities developed over the last few decades [10], ranging from mathematical models to those based on more intuitive approaches such as angles or distances. The most notable difficulties in aerial TD are discussed in [11] and include sensor noise effects, atmospheric attenuation and subsequent correction which can both lead to variabilities in target signature.

Depending on the system, remote sensing data can consist of high resolution RGB colour data, radar, multispectral or hyperspectral images. The latter, while providing a great deal of useful information, often at wavelengths beyond the range of human vision, introduces a vast quantity of data which must be handled and processed. Dimensionality Reduction (DR) techniques offer methods

of compressing and remapping this high dimensionality data into a reduced, and sometimes more informative, uncorrelated subspace. As hyperspectral images contain high levels of redundancy they are easily compressed using sparsity-based approaches [12] or by applying DR methods. Coupling spectral DR with TD in order to improve detection and classification rates has been covered widely in the literature [11,13–19] and has been shown to improve the performance of TD and classification algorithms.

In TD applications, often the targets are sparsely positioned in an imaged scene, therefore large amounts of spatial redundancy are often exhibited. This spatial redundancy, like the spectral redundancy also present in hyperspectral images, can be exploited in order to attain increased performance and efficiency. In [18,19], we investigated using the Normalised Difference Vegetation Index (NDVI) as a spatial mask on the detected image in order to constrain the region of interest in the scene. In this paper however, the spatial DR is applied prior to the calculation of the dimensionality reduced image in order to refine the subspace in which any TD is performed. NDVI and its variants are most often used in remote sensing applications to quickly and effectively assess vegetation health [3]. Other similar indices are used to detect water/snow in an image or for assessing how built upon an area is. However, such indices could be used to provide a measure of how informative a pixel may be or how likely it is to hold a target signature. Pixels are categorised as informative or non-informative with the non-informative pixels being discarded. By removing such pixels, the DR calculation can be simplified by reducing the number of samples, whilst also simplifying and suppressing the background class. As TD algorithms can be represented as a binary classification, improving the separation between target and background classes consequently improves TD performance [8]. While various information indices are commonly used in remote sensing tasks, to the best of the authors' knowledge, they have never been used to perform spatial DR or coupled with spectral DR in this way with the aim of improving hyperspectral TD applications.

In this paper, we investigate the use of coupled spatial and spectral DR on hyperspectral TD applications. With this approach, we aim to decrease both the spatial and spectral redundancy exhibited in hyperspectral images, improving the efficiency and performance of various benchmark TD algorithms. The proposed method was tested on two hyperspectral datasets containing multiple targets in varied scenes.

2. Materials and Method

In this section, we first introduce the notation used in this paper as well as the relevant background information on each of the datasets and methods used. Secondly the various spectral DR methods used are introduced followed by the spatial DR method created for purpose of TD. Finally the various detection algorithms are described.

2.1. Notation

Hyperspectral images can most easily be represented as 3-dimensional datacubes, with two spatial dimensions and a third spectral dimension. Any hyperspectral image **X** can be represented as L individual greyscale images each exposed at a particular wavelength or spectral band λ_l, $X_l : l \in \{1, 2, ..., L\}$, where L represents the total number of spectral bands. Alternatively an image, **X**, can be thought of as N individual pixels each comprised of an L-dimensional vectors as seen in Equation (1):

$$\mathbf{X}_{3D} = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{i,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{i,2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,j} & x_{2,j} & \cdots & x_{i,j} \end{bmatrix} \tag{1}$$

where i and j represent the number of columns and rows in the hyperspectral datacube $\mathbf{X}_{3D}$ respectively. Generally when applying hyperspectral image processing algorithms to images, it is desirable for the

75 image to be in a 2-dimensional matrix form, $\boldsymbol{X}$. This is shown in Equation (2), where each column
76 consists of a single pixel, $\boldsymbol{x}_i : i \in \{1, 2, ..., N\}$, represented by an L-dimensional vector, as seen in
77 Equation (3).

$$X_{2D} = \left[\begin{array}{ccccccc} \boldsymbol{x}_1, & \boldsymbol{x}_2, & \boldsymbol{x}_3, & \cdots & \boldsymbol{x}_i, & \cdots & \boldsymbol{x}_N \end{array} \right] \tag{2}$$

$$\boldsymbol{x} = \left[\begin{array}{cccc} x_{\lambda_1}; & x_{\lambda_2}; & \cdots & x_{\lambda_L} \end{array} \right] \tag{3}$$

78 The vector in Equation 3 represents a single hyperspectral pixel, or a single spectral measurement.

2.2. Image Acquisition

80 Images from two sources have been used to validate the techniques described here. The first
81 dataset "OP7", provided by BAE Systems, consists of three images acquired on the 18th May 2014
82 from an aerial platform flying at an altitude of approximately 0.78km. The platform used a Visible and
83 Near-InfraRed (VNIR) hyperspectral sensor with a spectral range of roughly 400 - 1000nm.
84 The second set of images were supplied by the UK Defence Science and Technology Laboratory
85 (DSTL) as part of the University Defense and Research Collaboration (UDRC) from the Selene trial.
86 Part of this trial collected airborne hyperspectral imagery of large numbers of spectrally varied targets
87 across a two week period between the 4th and 15th of August 2014 at an altitude between 0.9 and
88 1.05km. A common region from a selection of seven images captured over this period was used so as to
89 exhibit varied targets under different environmental conditions. The camera used was also in the VNIR
90 range with a similar spectral range of roughly 400nm to 1000nm with fewer spectral measurements
91 but a much higher spatial resolution than the OP7 dataset.
92 Sample false-colour images from each of the datasets can be seen in Figure 1 along with cropped
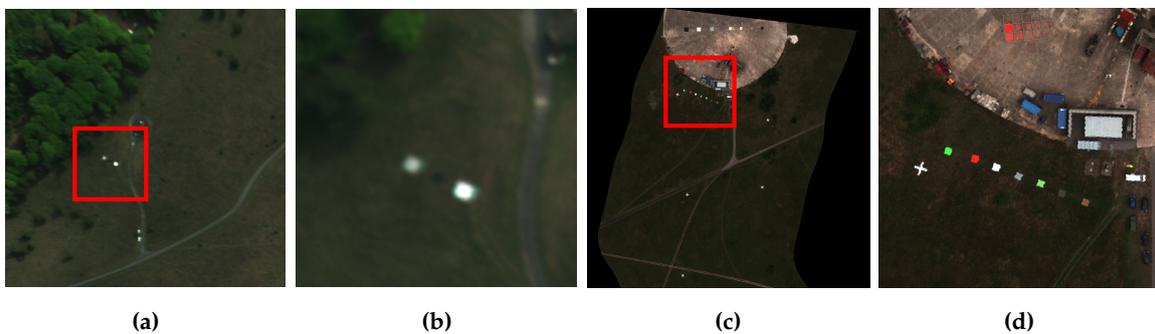93 portions of the target area indicated by a red box.

(a)   **(b)**   **(c)**   **(d)**

Figure 1. False colour images from the datasets used in this work. **a)** OP7 image. **b)** Target Region of a). **c)** UDRC Selene image. **d)** Target Region of c).

2.3. Spectral Dimensionality Reduction Techniques

95 Due to the high correlation between successive bands in hyperspectral images, compression and
96 DR techniques can be readily applied. In this section we review four of the most common techniques
97 which we have included in this analysis.

2.3.1. Principal Component Analysis

99 Principal Component Analysis (PCA) [20] is a classical method of DR. It seeks to remap highly
100 correlated data into an uncorrelated space using a set of optimal orthogonal basis vectors, or Principal
101 Components (PC), calculated from the input data. There are multiple ways of achieving this through
102 both iterative and non-iterative algorithms, we have included two in this analysis, Eigenvalue
103 Decomposition (EVD) and Non-linear Iterative Partial Least Squares (NIPALS). The EVD is a common

method for performing PCA and consists of the matrix decomposition $\Sigma = U\Lambda U^T$, where the matrix Λ is a diagonal matrix containing the eigenvalues of Σ, *i.e.*, $\Lambda = \text{diag}\{\lambda_1, \lambda_2, ..., \lambda_L\}$ and the matrix U contains the related eigenvectors $[u_1, u_2, ..., u_L]$. The eigenvalues in Λ are ordered such that $\lambda_1 > \lambda_2 > ... > \lambda_L$, hence the first K largest eigenvalues correspond to the first K eigenvectors. The first K eigenvectors, or PCs, can be used as the a set of basis vectors to transform the original data into an uncorrelated K-dimensional subspace, where $K < L$, which represents the most significant information contained in the data.

In some cases, such as those where the desired number of retained components is known, it is unnecessary and therefore preferable to avoid calculating every PC as is required in an EVD. In these cases iterative techniques can be used to calculate each successive PC in turn until the required number, K, has been reached. The NIPALS algorithm can be used to achieve this and consists of the decomposition $X = TP^T$, where X is some mean-centred matrix and the columns of T are the scores and the columns of P are the loadings. P forms an optimal transform matrix which can be used in an identical manner to the matrix of eigenvectors from an EVD in transforming input data into an dimensionality reduced subspace. An overview of the NIPALS algorithm can be found in [21].

In testing, both the EVD and NIPALS algorithms produced PCs with identical magnitudes but some which exhibited opposite polarity as orthogonality can take one of two directions. The EVD has no need to converge and is therefore faster while producing minimal error. With this in mind, only the EVD was used to perform PCA-based DR.

2.3.2. Maximum Noise Fraction

The Maximum Noise Fraction (MNF) [22] transform is similar in operation to PCA but also accounts for the noise present in input data [23]. Rather than ordering the PCs of an input image, X, by their variance, as in PCA, they are instead sorted by their estimated Signal-to-Noise Ratio (SNR). In MNF, it can be assumed that the covariance of the data, Σ, is a sum of the covariance of the signal, Σ_s, and the covariance of noise, Σ_n, *i.e.*, $\Sigma = \Sigma_s + \Sigma_n$. The MNF transform seeks to maximise the calculated eigenvalues with respect to the estimated SNR and can be interpreted as two separate PCAs computed in turn, the first to noise whiten the data, and the second to calculate the PCs. The complete MNF algorithm is described in [22].

2.3.3. Folded Principal Component Analysis

With both PCA and MNF, as well as many other PCA-like methods, it is necessary to compute the full covariance matrix Σ. This covariance matrix is of size $L \times L$ where L is equal to the number of spectral bands in an image. Therefore, for images with high spectral resolution it can be computationally expensive and time-consuming to compute. In order to circumvent this challenge, Folded Principal Component Analysis (FPCA) [24] seeks to reduce the size of the covariance matrix and also incorporate the correlation within spectra into the calculation. In order to perform FPCA, each of the N mean-centred spectral vectors, $\bar{x}$, are folded into an $H \times W$ matrix, A, where $H \times W = L$ for some positive integers H and W. A partial covariance matrix can be calculated as $\Sigma = A^T A$ and using each of these N partial covariance matrices the full covariance matrix, Σ_{FPCA}, can be calculated as $\Sigma_{FPCA} = \frac{1}{N} \sum_{i=1}^{N} A_i^T A_i$. Images can be projected into the FPCA domain by performing the EVD, of Σ_{FPCA} and using the resultant eigenvectors to project the input data into the PC space. Auxiliary target spectra can then be folded using the same H and W and projected using the eigenvectors of Σ_{FPCA}, before being unfolded again to be processed in the FPCA domain.

2.3.4. Independent Component Analysis

Independent Component Analysis (ICA) is a common method for performing Blind Source Separation (BSS) used in DR. Unlike PCA, MNF or FPCA, ICA seeks to separate an ensemble of mixed signals into a set of finite distinct sources or Independent Components (IC). This is achieved by maximising the statistical independence of the calculated components [25]. As hyperspectral images

are made up of a weighted sum of a set of finite pure spectra, or endmembers, it is possible to perform ICA to separate the mixed spectra into pure spectral endmembers. There are multiple algorithms used to calculate the ICs of a set of data, two of the most well used are the FastICA [26] algorithm and the Joint Approximation Diagonalization of Eigen-matrices (JADE) algorithm [27]. In this paper the FastICA algorithm is used instead of the JADE algorithm as it reached convergence both faster and more reliably. In order to perform ICA based DR, the required number of ICs to represent the data needs to be calculated. This is achieved by using the notion of Virtual Dimensionality (VD) [28] which estimates the number of spectrally distinct sources in the image. Using the method from [29], ICA-DR can be achieved with K ICs.

PCA and MNF are both classified as second order statistics-based transforms which can be insufficient in some applications [29]. ICA preserves higher order moments, such as skewness and kurtosis, which can aid in applications which require characterisation of subtle differences in signature such as classification or detection of small/rare targets. While it is possible that second-order statistics may be insufficient in preserving such characterising information it has not been the case with this application. Although it performs favourably when compared to other ICA algorithms such as JADE, FastICA is much slower than the other, non-iterative, methods for DR listed here. This is due to the need for multiple iterations to reach a convergence and is therefore another important consideration in its choice in any practical application.

2.4. Spatial Dimensionality Reduction using Vegetation Indices

As well as exploiting the spectral redundancy exhibited in hyperspectral images the spatial redundancy can also be utilised for TD through compression or by creating new features. By investigating the spectral properties of the scene, spatial areas of interest can be selected and areas of non-interest can be discarded from further processing, often saving on large computational costs. Vegetation Indices (VI) such as NDVI and its variants are of particular interest in TD applications as they offer simple and effective methods to discriminate between vegetative and non-vegetative pixels. Three NDVI variants were selected and tested in discriminating between the desired background of vegetation and the foreground of synthetic materials to which the target objects of interest belong. Each of the methods used in this work are listed in Table 1.

Table 1. Vegetation indices used for spatial DR

Index	Acronym	Equation	Reference
Normalised Difference Vegetation Index	NDVI	$\dfrac{\lambda_{NIR} - \lambda_{Red}}{\lambda_{NIR} + \lambda_{Red}}$	Rouse et al. [30]
Normalised Difference Vegetation Index (red-edge)	NDVI$_{re}$	$\dfrac{\lambda_{re} - \lambda_{Red}}{\lambda_{re} + \lambda_{Red}}$	Hansen & Schjoerring [31] Ettehadi et al. [2]
Red-Edge Normalised Difference Vegetation Index	RENDVI	$\dfrac{\lambda_{750} - \lambda_{705}}{\lambda_{750} + \lambda_{705}}$	Gitelson & Merzlyak [32] Sims & Gamon [33]

2.5. Target Detection Algorithms

In this paper, five common classical methods for TD and Anomaly Detection (AD) are investigated for use in combination with spatial and spectral DR. Four of these five algorithms, the Adaptive Cosine Estimator (ACE)[34], Constrained Energy Minimisation (CEM)[13], the Spectral Angle Mapper (SAM)[35], Spectral Information Divergence (SID)[36], and the Reed-Xiaoli Detector (RXD)[37]. Each method, with the exception of the latter are TD algorithms and, as such, they require *a priori* information about the targets to be detected in the form of a reference or ground truth spectra. The final method however, the RXD, does not require prior information about a target and finds outlying or anomalous pixels within the image and is cited as the benchmark AD algorithm [11]. Whilst other TD algorithms

such as Orthogonal Subspace Projection (OSP) [38,39] are often used to good effect [40–42], such methods require prior knowledge of the background which may not be fully known and as a result hinder the performance in a TD application hence they are left out of this analysis. ACE in particular has been shown to achieve favourable results in similar comparisons with other TD algorithms [11,14,17,43].

2.6. Performance Measures

In order to asses the performance of each of the TD algorithms a number of measures are used in this paper. Each of the various TD and AD algorithms used return a probability or confidence measure as to whether each pixel contains a target. By varying the threshold above which a pixel is classified as a target, the various behaviours and performance of a TD algorithm can be assessed. Both Receiver Operator Characteristic (ROC) curves [44] and Precision-Recall (PR) curves [45,46] are useful measures in determining an optimal operational threshold in order to maintain an acceptable False Alarm Rate (FAR). The Area Under the Curve (AUC) is a useful measure for comparing the ROC and PR behaviours of various algorithms. The ROC curve can be created by plotting the Probability of Detection (P_d), against the Probability of False Alarm (P_{fa}), at a series of thresholds.

Although ROC curves are a simple and effective way of rapidly visualising the performance of a classifier, it has been shown that ROC analysis can be flawed for unbalanced classes, as is the case for TD applications. In [45] it is shown that PR curves are more informative for unbalanced classes as they correctly evaluate the fraction of True Positive (TP) detections amongst the total number of positive predictions, or the precision of the classifier. Precision can be calculated using the number of TP and False Positive (FP) detections where precision $= TP/(TP + FP)$. Recall is calculated using the number of False Negative (FN) detections where recall $= TP/(TP + FN)$. Recall is the fraction of TP detections amongst the total number of positive examples. In the case of TD applications the number of positive examples is the total number of target pixels present in an image and PR curves can be obtained by plotting the precision of a classifier against its recall at a series of thresholds.

Along with these graph-based methods, four other methods of assessing each of the TD algorithms were used. Three measures commonly used in assessing binary classifier performance, the F1 score [46], Matthew's Correlation Coefficient (MCC) [47] and balanced accuracy [48] were used. As TD algorithms can be represented as a binary classification between a positive target class and a negative background class, these measures are applied to assess how each algorithm performs. The final metric used in this work is the visibility measure [14]. Visibility is an indication of how distinct a target is from its background. This is useful in assessing how the detection can be affected by applying DR to input image data.

2.7. Proposed Methodology

In this paper, we are proposing a pipeline to improve TD in hyperspectral images by combining both spatial and spectral DR methods. This is achieved by performing a spatial DR on an input image, removing any vegetative, and therefore, non-target pixels, before projecting the subset of the image into a subspace using more traditional spectral dimensionality methods. Any relevant ground truth target spectra can also be projected into the same subspace using the forward transform of each DR method. The TD can be performed in the dimensionality reduced subspace. This pipeline is shown in Figure 2.

In previous work, [18,19], both NDVI and PCA were combined to improve the performance of a hyperspectral Hit-or-Miss Transform (HMT) for use as a TD algorithm. By reducing the spatial and spectral redundancy the computational overhead of the proposed algorithm was reduced. NDVI was used to mask the already dimensionality reduced data. However, this meant that the NDVI had no influence over the performance of the detection. When it is already known that the target is non-vegetative, the application of Normalised Difference Vegetation Index (red-edge) ($NDVI_{re}$) masking prior to the use of spectral DR improves the performance of TD algorithms because a
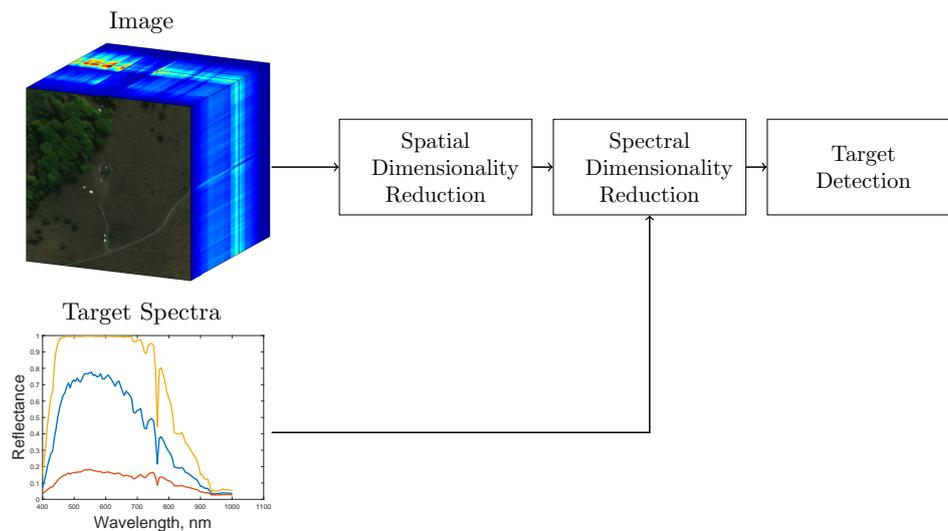
Figure 2. Spatio-spectral dimensionality reduction pipeline for target detection.

much more informative subset of pixels is exploited. Rather than using the spectral information of vegetation in the DR calculation, which can skew the resultant basis vectors away from representing desirable signatures, it is instead overlooked. The DR is targeted towards representing potentially more informative pixels. By suppressing the vegetative part of the background class, an improved separation between the target and remaining background can be achieved in the DR subspace. The aim is, that by reducing the number of samples in this way, the calculation of the dimensionality reduced data is not only simplified but also more useful information is retained in potentially fewer components.

3. Experimental Results

In this section we will first investigate the achievable compression rates when combining both spectral DR and $NDVI_{re}$-based spatial DR. We then select the optimal detection algorithm for use with the proposed spatio-spectral DR pipeline shown in Figure 2. Then, we present a subset of the results gathered using both the OP7 dataset and the UDRC Selene Trial data. Finally, we investigate the effects of the various spatial and spectral DR schemes combined with the chosen TD algorithm. Each of the spectral DR methods, PCA, MNF, FPCA and ICA, are tested with $K = 20$ components retained and "Raw" refers to the full dimensionality image where $L = 100$ for the OP7 data and $L = 80$ for the Selene trial images.

3.1. Selection of the optimal Vegetation Index for Spatial Dimensionality Reduction

In order to assess which VI gave the best separation between vegetative and non-vegetative pixels, the ground truth spectra of multiple green targets from the Selene dataset as well as the average spectra of a patch of vegetation were investigated. Figure 3 shows the test image used as well as the results of each of the three VIs.

All three of the VIs are able to identify a good separation between vegetation and most other non-vegetative background pixels. However some of the green targets present in the scene, despite exhibiting distinctly non-vegetative spectra, can produce a ratio similar to that of the surrounding grass, this is most apparent when using the basic NDVI. The regions investigated are indicated by the blue and orange elements in Figure 3, matching the colour of the plotted spectral signatures in Figure 4. Figure 4 shows the target spectra, background spectrum, and VI bands used to calculate the ratio of each VI result respectively.

Two of the targets from Figure 3a, green perspex (circled in blue) and green ceramic (circled in orange) were investigated for separation from the background when using VI-based spatial DR. Each
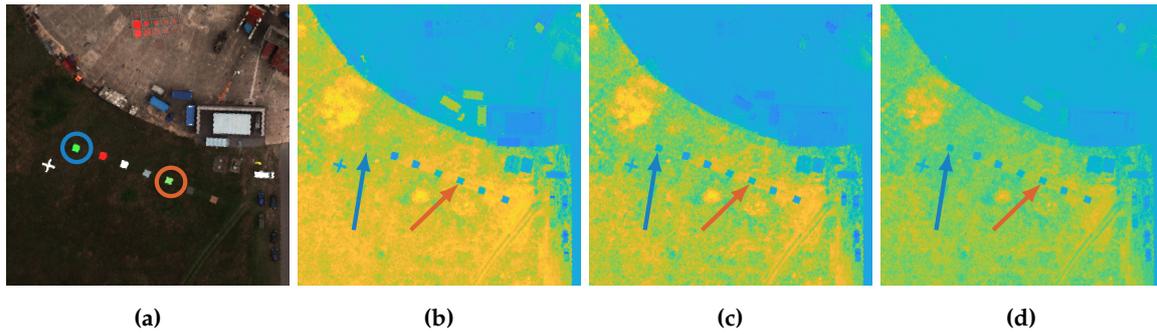
(a)   **(b)**   **(c)**   **(d)**

Figure 3. Optimal VI experiments **a)** UDRC Test Image. **b)** NDVI ratio. **c)** NDVI$_{re}$ ratio. **d)** RENDVI ratio. (In **b)** to **d)** warmer colours indicate higher levels of vegetation and colder colours indicate non-vegetation.)
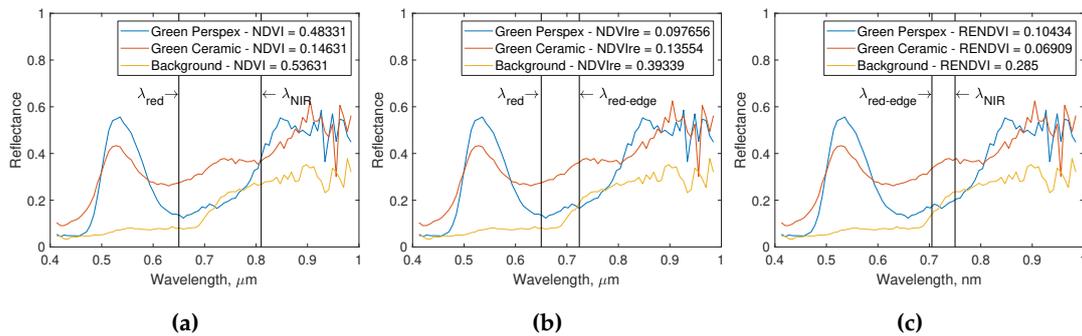

(a)   **(b)**   **(c)**

Figure 4. VI ratio of each of the three test spectra **a)** NDVI ratio and band locations. **b)** NDVI$_{re}$ ratio and band locations. **c)** RENDVI ratio and band locations.

of the three VIs investigated produce a ratio between the intensity of a pixel at two bands, the two targets produced VI values shown in Table 2

Table 2. Vegetation index ratios obtained for background and targets

Vegetation Index	Green Perspex Ratio	Green Carpet Ratio	Background Ratio
NDVI	0.48	0.14	0.53
NDVI$_{re}$	0.09	0.13	0.39
RENDVI	0.10	0.06	0.28

From Figures 3b-3d, Figure 4 and Table 2, it is possible to see that NDVI and Red Edge Normalised Difference Vegetation Index (RENDVI) have lower separability between the "green perspex" target and the background. In fact, it can be observed that the green perspex target, pinpointed by the blue arrows in Figures 3b-3d, is near indistinguishable from the background in Figure 3b with only six of the seven targets having a low enough NDVI value to be reliably distinguished from the background. Despite having a distinct spectral profile, as shown in Figure 4a, the green perspex has an almost identical NDVI value to the background (0.48 vs 0.53) indicating the ratio between the two NDVI bands is nearly the same. By altering the Near-InfraRed (NIR) band to be placed in the red-edge portion of the spectrum, as is the case when using NDVI$_{re}$, a much greater separation is achieved (0.09 vs 0.39). This is due to the red-edge phenomenon, when the intensity of the background spectra rises sharply, reflecting NIR light. RENDVI, whilst successfully segmenting all seven targets in this example, creates a lower contrast between background and target when compared with NDVI$_{re}$. As NDVI$_{re}$ provides the best separation between the most difficult targets and the background it is used to implement spatial DR in this paper.

3.2. Combining spatial and spectral DR for hyperspectral compression

Here we briefly discuss the effects on image size and compression when combining spatial and spectral DR techniques. $NDVI_{re}$ is used as a spatial mask, selecting pixels that are relevant and can be used in subsequent spectral DR and TD processes. By masking certain pixels they can be discarded from further processing, reducing the sample size. Then, by performing spectral DR, retaining K components from L spectral bands the sample size is reduced further. By combining the remaining spatial and spectral components, a compressed representation of the relevant data is retained for further processing. Table 3 details the size of each of the images used in this paper, as well as their compressed spatial and spectral sizes along with the percentage of the original data retained after compression.

Table 3. Achieved compression for combined spatial and spectral DR

Image	# Samples Full	# Samples $NDVI_{re}$	L	K	Spatial Comp.(%)	Spectral Comp.(%)	Total Comp. (%)	Average Comp. (%)
OP7_1	160000	3504	100	20	2.19	20	0.44	
OP7_2	160000	2500	100	20	1.56	20	0.31	0.34
OP7_3	160000	2232	100	20	1.40	20	0.28	
IM_140804	3210191	649435	80	20	20.23	25	5.06	
IM_140806	3839976	578674	80	20	15.07	25	3.77	
IM_140807	3415052	689245	80	20	20.18	25	5.05	
IM_140808	3015944	543569	80	20	18.02	25	4.51	4.61
IM_140812	4360159	610172	80	20	13.99	25	3.50	
IM_140813	3301404	807262	80	20	24.45	25	6.11	
IM_140815	3640769	626776	80	20	17.22	25	4.30	

The OP7 dataset images are first able to be compressed to 1.72% of their original size on average as $NDVI_{re}$ selects a small proportion of the total pixels to process further. By retaining $K = 20$ components in the subsequent spectral DR stage, this is reduced further to an average of 0.34% of their original size. The images in the Selene trial have a much higher spatial resolution and a larger sample is retained after using the $NDVI_{re}$ spatial mask as a large proportion of the pixels represent non-target and non-vegetative materials, as shown in Figure 1. The pixels retained after $NDVI_{re}$ represent an average of 18.45% of the original image and applying spectral DR, with $K = 25$, reduces this to 4.61% on average.

3.3. Comparison of the TD Algorithms used

Each of the detection algorithms used were individually tested for their suitability when combined with the spatial and spectral DR schemes selected. In order to validate which algorithm performed optimally, the proposed method was applied to a subset of the Selene data. First an ROC analysis was performed with examples of ROC curves for each combination of TD and DR algorithms are shown in Figure 5 for the full spatial scene and in Figure 6 when combined with $NDVI_{re}$. Figures 5 and 6 show the upper left quadrant of the ROC curves in order to highlight the differences between each of the methods used.

As previously stated, ROC analysis, in isolation, is insufficient for comparing unbalanced binary classifiers [45]. However, it is interesting to note that the disparity between the ROC curves from each of the TD algorithm outputs. In Figure 5 each of the algorithms used have near ideal ROC curves regardless of which spectral DR scheme is used when working on the full spatial scene. However when spatial DR is employed, only the ACE and CEM algorithms remain near ideal as seen in Figure 6. The AUC of the ROC curves increase for each spectral DR scheme when combined with $NDVI_{re}$-based spatial DR and the ACE algorithm, as shown in Figure 6a. By simplifying the background, and therefore improving the covariance estimate, the ACE algorithm can achieve better separation between
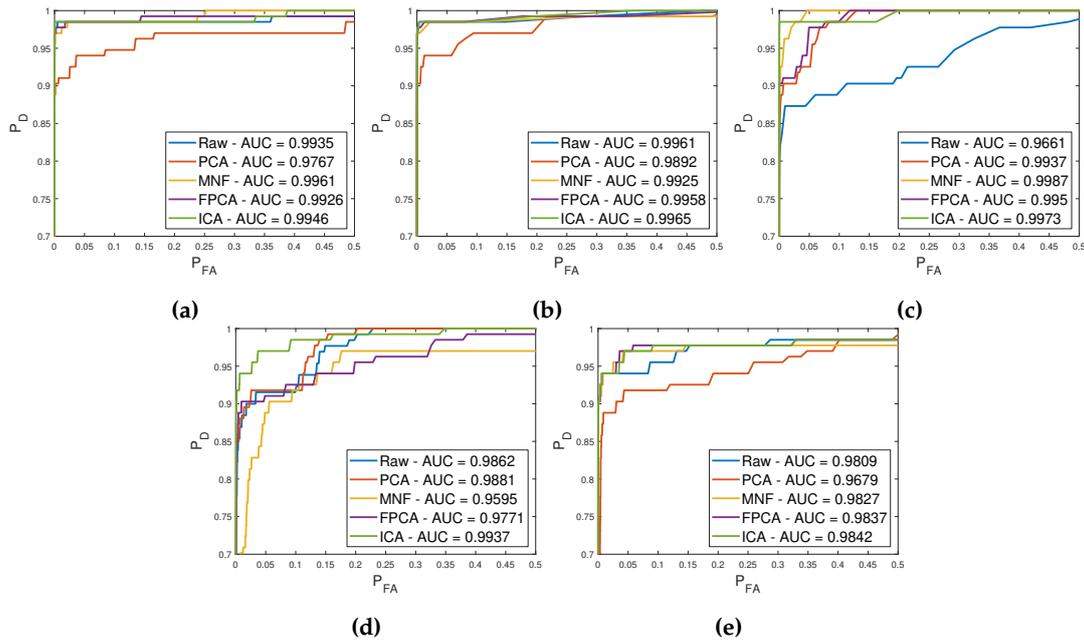
Figure 5. ROC Curves for each TD and spectral DR scheme pairing on the full scene. **a)** ACE. **b)** CEM. **c)** SAM. **d)** SID. **e)** RXD.

the known target and the estimated background. Similarly, by suppressing the background, the FIR filter weight estimation that is necessary for the CEM algorithm is simplified. This is reflected in the increased AUC values of the ROC curve when using CEM with $NDVI_{re}$-based spatial DR, as shown in Figure 6b.

Figure 6. ROC Curves for each TD and spectral DR scheme pairing in combination with spatial DR. **a)** ACE. **b)** CEM. **c)** SAM. **d)** SID. **e)** RXD.

As well as ROC curves, PR curves were generated for each of the combinations of TD and DR algorithms with and without the $NDVI_{re}$ spatial DR. The PR curves of each of the TD algorithms when considering both the full spatial scene and with the application of $NDVI_{re}$-based spatial DR are shown in Figure 7.
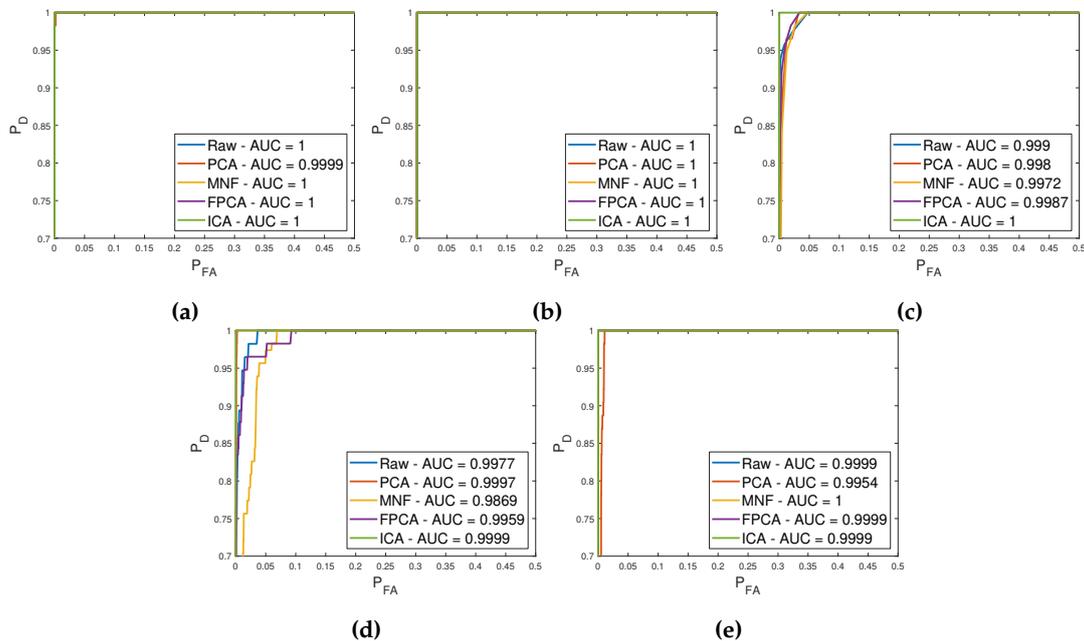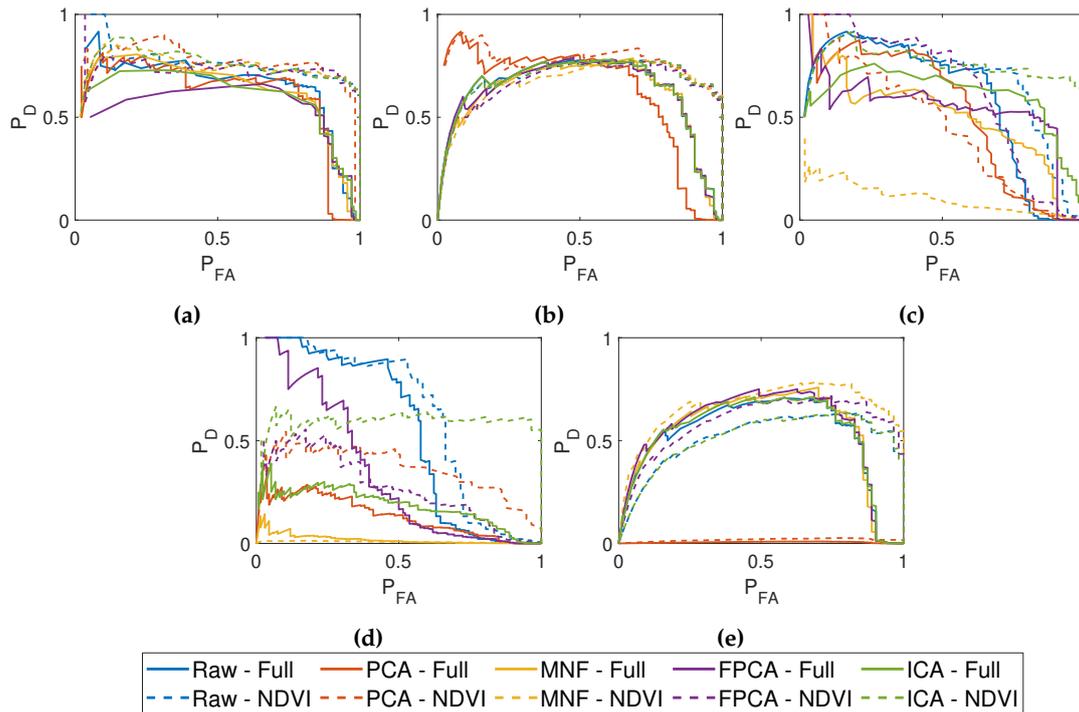
Figure 7. PR Curves for each TD and spectral DR scheme pairing in combination with spatial DR. **a)** ACE. **b)** CEM. **c)** SAM. **d)** SID. **e)** RXD.

Table 4. Comparison between the AUC of the PR curves using the full and spatial DR images

PR	Raw		PCA		MNF		FPCA		ICA	
AUC	Full	NDVI$_{re}$	Full	NDVI$_{re}$	Full	NDVI$_{re}$	Full	NDVI$_{re}$	Full	NDVI$_{re}$
ACE	0.649	**0.7556**	0.6038	**0.7505**	0.6229	**0.7393**	0.56	**0.7186**	0.5999	**0.7507**
CEM	0.6207	**0.6852**	0.6208	**0.7673**	0.6033	**0.6633**	0.6124	**0.6669**	0.6195	**0.6849**
SAM	0.577	**0.6723**	**0.5127**	0.4443	**0.4938**	0.0993	0.528	**0.6194**	0.6006	**0.7507**
SID	0.5315	**0.6112**	0.131	**0.3582**	0.0187	0.0102	**0.3314**	0.2625	0.1809	**0.5871**
RXD	**0.5153**	0.5086	0.0055	**0.0175**	0.5358	**0.6604**	0.5445	**0.5816**	**0.5224**	0.5049

Investigating the PR curves shown in Figure 7 and the corresponding AUC values in Table 4 we see that using ACE, CEM and SAM generally all yield high AUC values for each of the spectral DR schemes used. When NDVI$_{re}$-based spatial DR is used in combined with the spectral DR the AUC in almost every case, including on the raw data where no spectral DR is used. SID, when used on the full dimensionality data, provides an average AUC which is once again improved when using NDVI$_{re}$-based spatial DR. The RXD performs well when using the full data and each of the spectral DR algorithms with the exception of PCA where it fails to discriminate target materials. This is due to the fact that, mathematically, PCA is the inverse operation of the RXD [49]. PCA exploits the redundancy of hyperspectral images by only retaining the PCs corresponding to the largest eigenvalues whereas the RXD works by investigating the anomalous data attributed to smaller eigenvalues which have been discarded.

Both the ROC and PR analysis were performed on a per-target basis. The results shown in Figures 5-7 and Table 4 are from the detection of a single target however they are generally representative of the performance over every target present in the scene. Along with the ROC and PR curves, the other performance measures detailed in Section 2.6 were calculated for each of the targets in the scene. These measures were then averaged in order to obtain an overview of each TD algorithm's general performance, the results of which can be seen below in Figure 8.
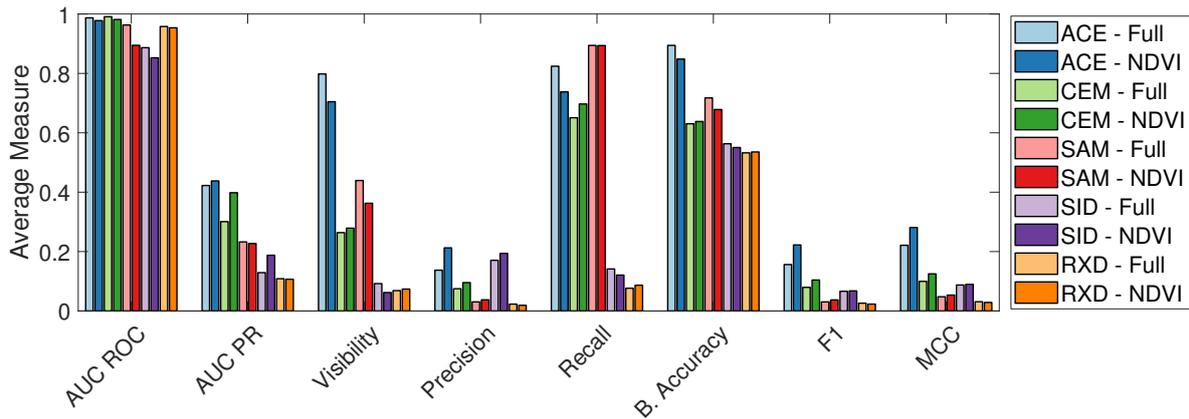
Figure 8. Average performance of each TD and DR algorithm combination when used on the full scene vs. when combined with spatial DR.

Similar to the results drawn from Figures 2 and 5, each of the TD algorithms perform well when considering the AUC of the ROC curves. ACE and CEM give the highest AUC of the PR curves with ACE, CEM, and SID each performing better when combined with spatial DR. Generally using the spatial DR reduces the visibility with the exception of CEM and the RXD where it slightly increases. ACE gives the highest visibility when considering both the full scene and when using spatial DR indicating it is the best at separating the background from the target of the algorithms investigated. ACE and SID display the best precision, with both methods improving when using spatial DR. ACE also displays the highest balanced accuracy, F1 score and MCC of each of the detectors tested. For these reasons, the remaining results in this paper are generated using the ACE algorithm solely. It is interesting to note that, as well as reducing the sample size for increased efficiency, as seen in Figure 8, for each of the detection algorithms, the performance after the application of spatial DR is generally as good or an improvement over using the full scene.

3.4. Results on the OP7 Dataset

The first of two datasets used in this paper was provided by BAE System. It consists of three images of a forest scene each portraying a common target area from overlapping views. The target area contains three calibration panels, one grey, one black and one white, which were used as the targets of interest. Figure 9a shows a false colour representation of one of the images with all three targets present in the scene. Figure 9c shows same the image masked using the NDVI$_{re}$ method detailed in Section 2.4. Figures 9b and 9d are enlarged views of the target areas of Figures 9a and 9c respectively.

| (a) | (b) | (c) | (d) |

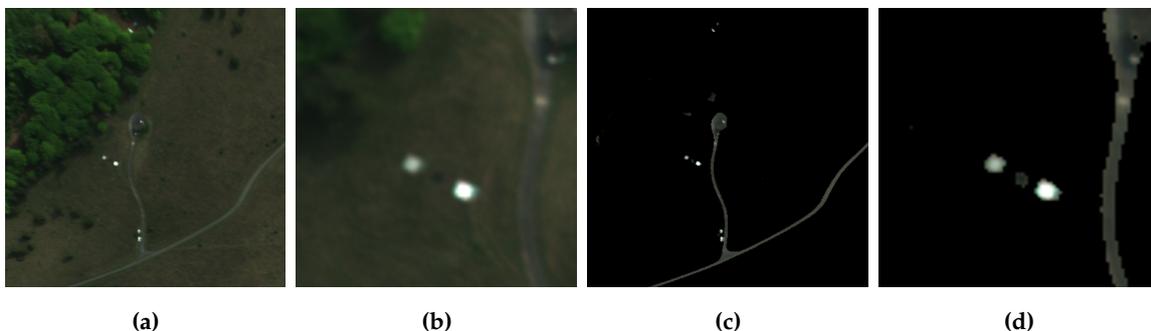**Figure 9.** Example of the OP7 Dataset. **a)** False colour image of the target area. **b)** Enlarged version of a).**c)** Retained pixels following the NDVI$_{re}$ spatial masking. **d)** Enlarged version of c).

Of the two datasets, OP7 is simpler as it contains fewer distinct materials and objects in the scene compared to the images from the Selene Trial dataset. The OP7 images also have a lower
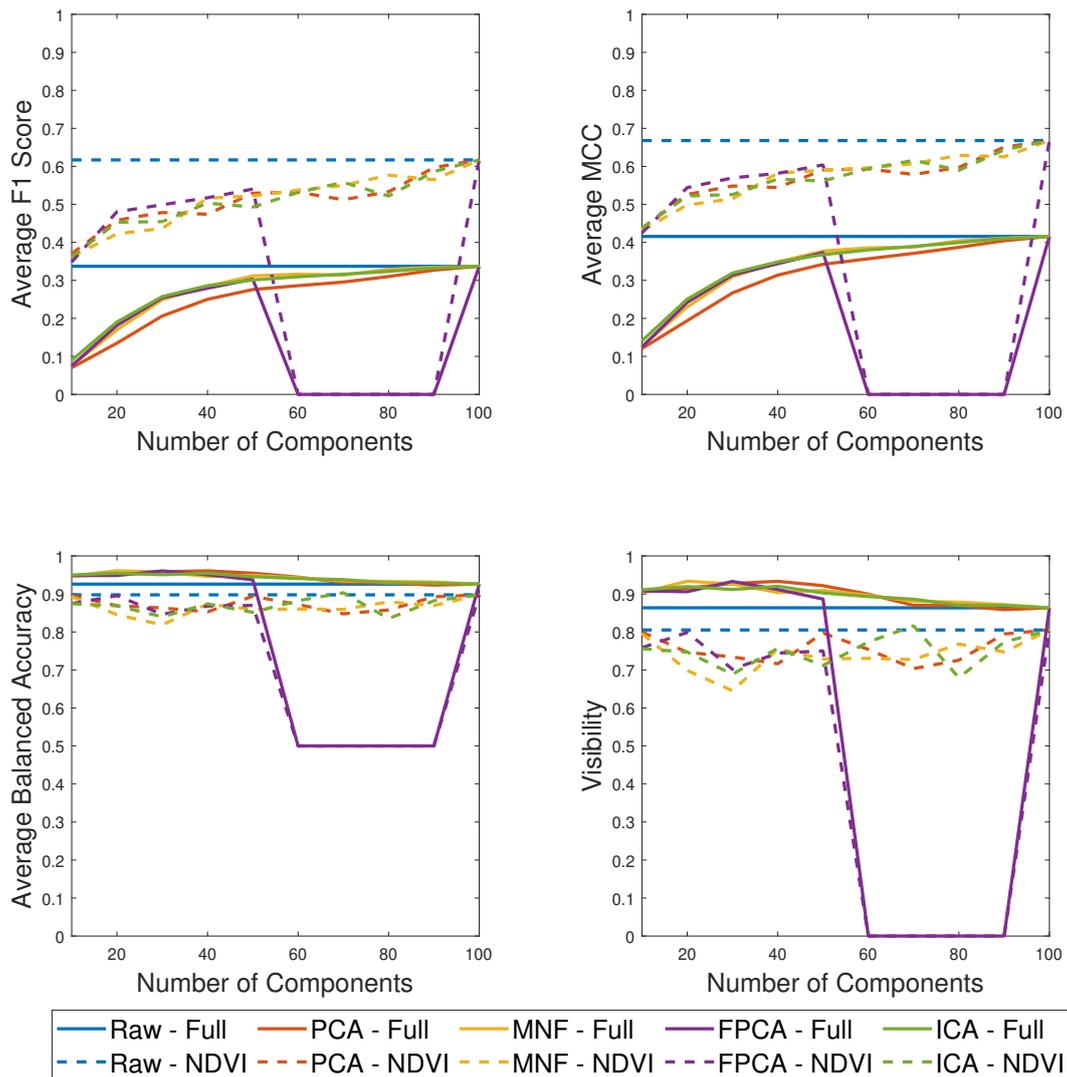
Figure 10. Performance measures using the ACE TD algorithm and each DR scheme with a varying number of retained components on the OP7 dataset.

363 spatial resolution when compared to the Selene Trial data, with a Ground Sample Distance (GSD) of
364 approximately 1 metre. As a result roughly nine pixels per target contain pure spectra.
365 In order to assess how each TD algorithm's behaviour varied with the number of components
366 retained using each DR scheme, the F1 score, MCC, balanced accuracy and visibility were calculated at
367 various values for K between $K = 10$ and $K = L$, where $L = 100$ for the OP7 data as shown in Figure
368 10. It must be noted that, when using FPCA, the dip in performance in each measure is a consequence
369 of an implementation limitations which result in the creation of a singular matrix. This restricts the
370 choice of the number of retained components and is discussed further in Section 4.
371 As seen in Figure 10, both balanced accuracy and visibility are largely invariant to the number of
372 components, K, retained. Although interestingly, at lower values of K, the visibility using each spectral
373 DR methods is greater than that of the raw data. Conversely, the F1 score and MCC both vary as the
374 number of components increase to be equal to the original spectral dimensionality of the data, both
375 with and without the application of spatial DR. This is to be expected, in the case where $K = L$ the data
376 is functionally identical, although it has been remapped, and no information has been lost in the DR
377 operation. By using spatial DR prior to spectral DR, both the F1 score and MCC are increased above
378 what is achieved using the raw full dimensionality data without spatial DR.

379   When comparing TD performance on the full spatial dimensionality images with that of the
380   NDVI$_{re}$ masked images, both with and without the application of spectral DR, the F1 score and MCC
381   both increase. However when spatial DR is applied, the average F1 score and MCC are considerably
382   higher. The removal of the vegetative background discards highly disparate observations and simplifies
383   the problem of separating background from target. This increases the precision of the detection as seen
384   in Tables 5 and 6. By reducing the complexity of the background, the targets, although more similar to
385   the remaining background, can be separated in the subspace more easily.

386   The MCC, in comparison to the F1 score, is slightly higher in both spatial DR cases as it takes
387   into account the correct identification of the true negative class. The balanced accuracy drops slightly
388   wherever NDVI$_{re}$ is applied. As the balanced accuracy is the average of the True Positive Rate (TPR)
389   and True Negative Rate (TNR), the decrease in the size of the True Negative (TN) class, without a
390   corresponding proportional decrease in FP, results in a lower balanced accuracy. Despite the increase
391   in False Positive Rate (FPR) when using NDVI$_{re}$, the absolute number of FP detections decreases. It can
392   also be seen in Figure 10 that, by removing the easily separated vegetative background using NDVI$_{re}$,
393   the visibility of the targets decreases. This occurs because the materials which remain are on average
394   more similar to the targets.

395   Further comparisons were made by retaining 20 components from each of the spectral DR methods
396   as this provided a good balance of performance and compression. As shown in Figure 10, $K = 20$
397   components also gave clear improvements over the raw, full dimensionality, scene when combined
398   with spatial DR. The improvement in detection when using spatial DR can be seen in Figure 11 where
399   there is less confusion in the detection map where NDVI$_{re}$ is applied, Figure 11b. The target is the
400   brightest object in the scene in each case, indicating good separability from the background.
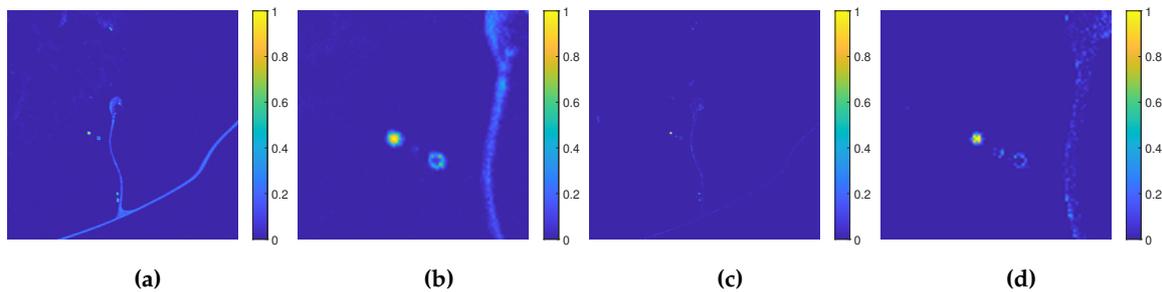

Figure 11. Detection statistics of the ACE algorithm on the full dimensionality data where yellow indicates a response of 1 and blue represents a response of 0. **a)** Without spatial DR. **b)** Enlarged version of a). **c)** With NDVI$_{re}$. **d)** Enlarged version of c).

401   In order to quantify this improvement, the ROC and PR curves for both the full and spatial
402   dimensionality reduced images are shown in Figure 12 for each spectral DR method where $K = 20$.

403   The ROC curves in Figures 12a and 12b are of the ACE detection statistics on the full scene and
404   NDVI$_{re}$ masked scene respectively. The AUC of the ROC curves alone is not significant as, regardless
405   of spatial and spectral DR used, it remains nearly identical. The AUC of the PR curves (Figures 12c
406   and 12d) when using the raw uncompressed data, PCA or MNF dimensionality reduced data increases
407   when spatial DR is applied. However, when applying FPCA the AUC falls slightly and when using
408   ICA the AUC of the PR curve decreases by 10%.

409   The results from Figures 10-12 are all calculated from a single target in order to display an example
410   of the performance achieved. The average results for each target are shown in Table 5 when considering
411   the full scene and in Table 6 when spatial DR has been applied.

412   In general, as shown in Tables 5 and 6, the AUC of the ROC curves are similar regardless of
413   the spectral and spatial DR used. The AUC of the PR curves varies with the spectral DR used with
414   each of the methods providing an average AUC. On average, employing spectral DR maintains the
415   performance when considering the full spatial scene but when combined with spatial DR there is a
416   slight decrease in the AUC of the PR curves. Applying employing NDVI$_{re}$-based spatial DR improves
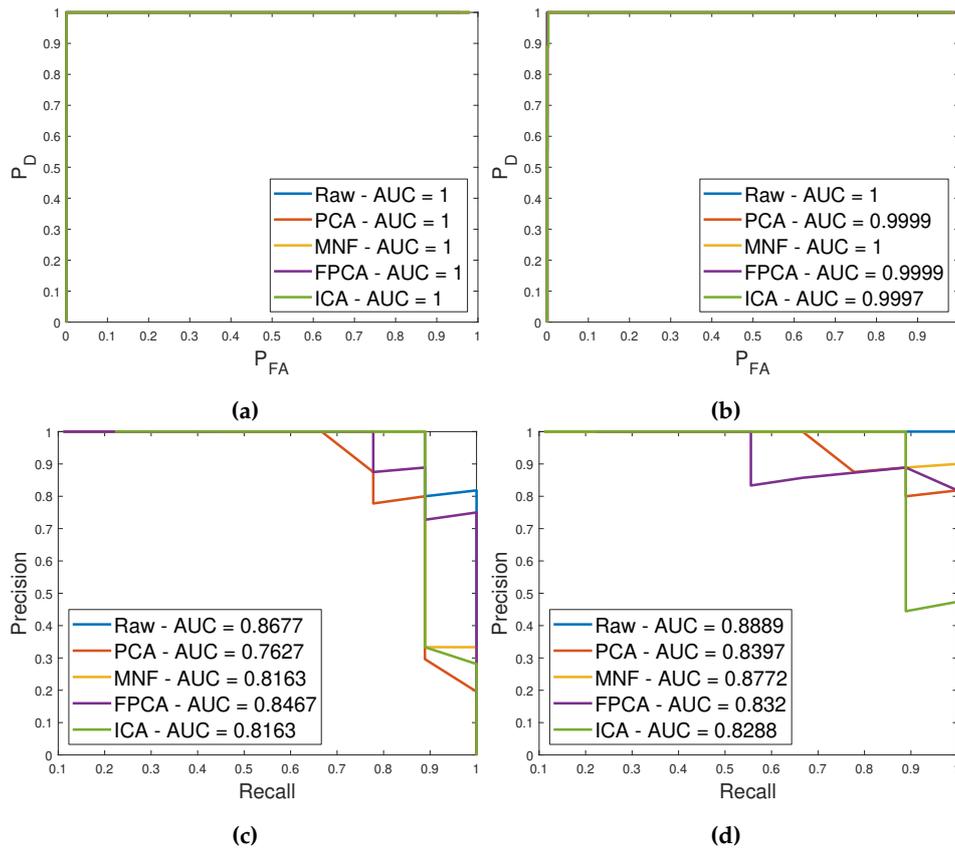
Figure 12. ROC and PR curves for detecting Target 1 (grey tile) in the OP7 dataset **a)** ROC curve using the full spatial dimensionality data. **b)** ROC curve using spatial DR pre-processing. **c)** PR curve using the full spatial dimensionality data. **d)** PR curve using spatial DR pre-processing.

Table 5. Average performance measures for each target in the OP7 dataset without spatial DR applied using the ACE algorithm.

K = 20	DR	ACE-Full							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
	Raw	**1.00**	**0.84**	0.88	**0.37**	0.88	0.93	**0.35**	**0.43**
	PCA	**1.00**	0.77	**0.94**	0.12	**0.95**	0.96	0.12	0.18
Grey Tile	MNF	**1.00**	0.79	0.93	0.15	**0.95**	0.96	0.15	0.21
	FPCA	**1.00**	0.80	0.92	0.17	0.94	**0.96**	0.17	0.23
	ICA	**1.00**	0.78	0.93	0.16	**0.95**	0.96	0.16	0.23
	Raw	**1.00**	0.06	0.60	**0.09**	0.62	0.80	0.06	0.11
	PCA	**1.00**	0.10	0.68	0.05	0.72	0.84	0.05	0.09
Black Tile	MNF	**1.00**	**0.15**	0.70	**0.09**	0.75	**0.85**	**0.07**	**0.12**
	FPCA	**1.00**	0.13	**0.71**	0.05	**0.76**	0.85	0.05	0.09
	ICA	**1.00**	0.11	0.67	0.04	0.72	0.83	0.05	0.09
	Raw	**1.00**	**0.74**	0.79	**0.57**	0.79	0.89	**0.53**	**0.59**
	PCA	**1.00**	0.67	**0.93**	0.22	**0.94**	0.96	0.28	0.37
White Tile	MNF	**1.00**	0.68	0.85	0.39	0.86	0.92	0.41	0.47
	FPCA	**1.00**	0.60	0.83	0.33	0.84	0.91	0.35	0.41
	ICA	**1.00**	0.67	0.79	0.47	0.80	0.89	0.44	0.49
	Raw	**1.00**	**0.55**	0.76	**0.34**	0.77	0.87	**0.32**	**0.37**
	PCA	**1.00**	0.52	**0.85**	0.13	**0.87**	0.92	0.15	0.22
All Spectra	MNF	**1.00**	0.54	0.83	0.21	0.85	0.91	0.21	0.27
	FPCA	**1.00**	0.51	0.82	0.18	0.85	0.91	0.19	0.25
	ICA	**1.00**	0.52	0.80	0.22	0.82	0.89	0.22	0.27

Table 6. Average performance measures for each target in the OP7 dataset with NDVI$_{re}$-based spatial DR applied using the ACE algorithm.

K = 20	DR	ACE-NDVI$_{re}$							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
Grey Tile	Raw	**1.00**	**0.86**	0.72	**0.73**	0.74	0.86	**0.59**	**0.65**
	PCA	**1.00**	0.83	0.75	0.56	0.79	0.87	0.48	0.54
	MNF	**1.00**	0.85	0.76	0.56	0.80	0.87	0.48	0.54
	FPCA	**1.00**	0.84	**0.81**	0.54	**0.84**	**0.90**	0.51	0.57
	ICA	**1.00**	0.75	0.73	0.52	0.77	0.86	0.43	0.50
Black Tile	Raw	**0.98**	**0.37**	**0.52**	**0.40**	**0.57**	**0.76**	**0.25**	**0.33**
	PCA	0.94	0.08	0.48	0.06	0.56	0.74	0.09	0.14
	MNF	0.96	0.09	0.46	0.07	0.54	0.73	0.09	0.14
	FPCA	0.94	0.09	0.47	0.08	0.54	0.73	0.10	0.15
	ICA	0.93	0.08	0.42	0.05	0.49	0.71	0.08	0.12
White Tile	Raw	**0.97**	**0.66**	0.57	**0.83**	0.58	0.78	0.58	**0.63**
	PCA	0.95	0.61	0.59	0.78	0.60	0.79	**0.59**	**0.63**
	MNF	0.95	0.62	**0.61**	0.74	**0.62**	**0.80**	0.58	0.62
	FPCA	0.92	0.59	0.59	0.64	0.61	0.79	0.51	0.55
	ICA	0.94	0.63	0.57	0.73	0.59	0.78	0.53	0.58
All Spectra	Raw	**0.98**	**0.63**	0.61	**0.65**	0.63	0.80	**0.47**	**0.53**
	PCA	0.96	0.50	0.61	0.47	0.65	0.80	0.39	0.44
	MNF	0.97	0.52	0.61	0.46	0.65	0.80	0.38	0.44
	FPCA	0.96	0.50	**0.62**	0.42	**0.66**	**0.81**	0.37	0.43
	ICA	0.96	0.49	0.57	0.43	0.62	0.78	0.35	0.40

417 the AUC when considering the full dimensionality data. The precision of the spatial DR coupled
418 methods is increased in comparison to using the full spatial scene as certain false positives are removed
419 either directly via the masking operation or indirectly by improving the spectral DR calculation. The
420 recall drops slightly, however this may not be significant in TD applications as one pixel on target is
421 sufficient for the identification and classification of an object of interest. Figure 10 shows a drop in the
422 visibility and balanced accuracy measures when applying spatial DR which is which is consistent for
423 each of the targets, as shown in Figure 13.

424 The visibility drops significantly when using the spatial DR as the highly dissimilar vegetative
425 background is removed making the average background and target spectra more similar. The balanced
426 accuracy falls when using NDVI$_{re}$ as the TN class decreases without a corresponding drop in FP
427 detections. The F1 score and MCC both increase when using NDVI$_{re}$-based spatial DR when applied to
428 the full dimensionality data as well as for each spectral DR scheme used. In nearly all of the measures
429 tested, the full spectral dimensionality image with and without spatial DR performed the best of all
430 methods on average with the application of spatial DR tending to improve the performance. Each
431 of the spectral DR methods employed retained only 20 components of the original 100, reducing the
432 computational complexity and cost of performing the TD while maintaining similar performance.
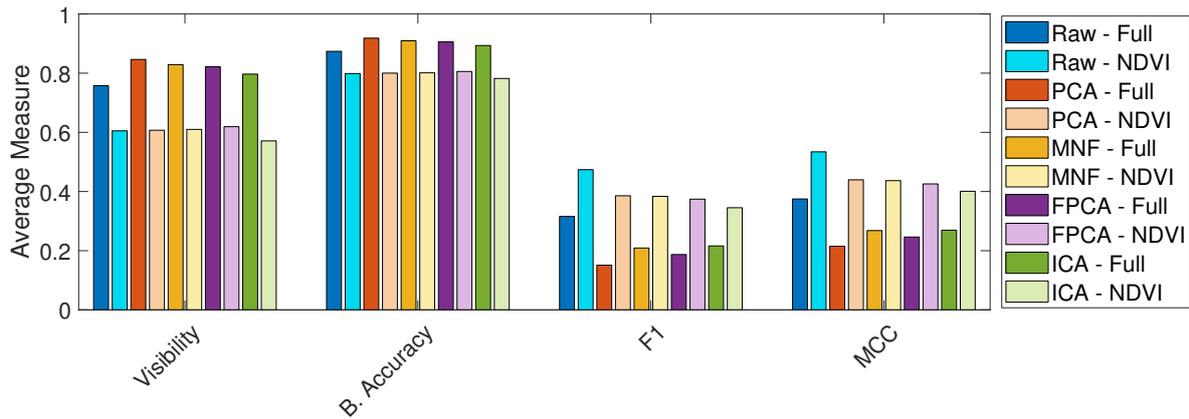
Figure 13. Comparison between the performance measures when combining spatial and spectral DR on the OP7 dataset

3.5. Results on the UDRC Selene Dataset

The second of the two datasets used in this paper was provided by DSTL. It consists of seven images of a different forest scene with a large concrete area with metal containers, vehicles and other objects captured over the course of two weeks in August 2014. Each image covers a different view of this common target area containing between five and seven calibration panels of various colours and materials with a GSD of roughly 0.3m.

Figure 14c shows the image masked using the $NDVI_{re}$ method detailed in Section 2.4 with Figures 14b and 14d providing an enlarged view of the target area from the Figures 14a and 14c respectively.

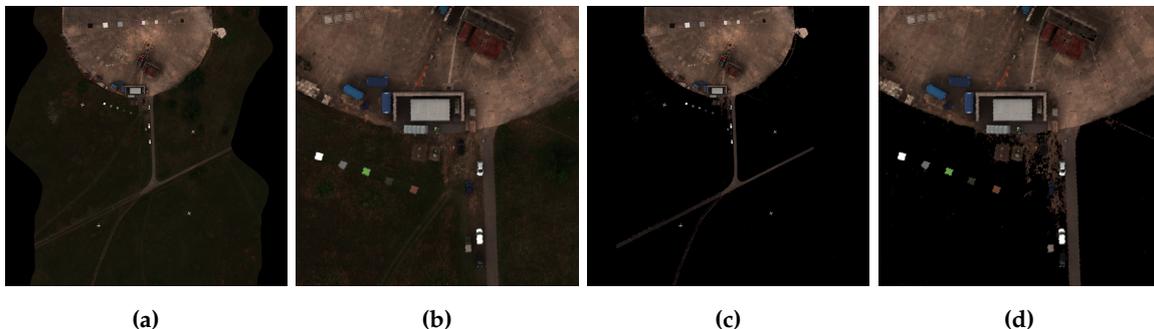

| (a) | (b) | (c) | (d) |

Figure 14. Example image from the UDRC Selene Dataset. **a)** False colour image of the target area. **b)** Enlarged version of a).**c)** Retained pixels following the $NDVI_{re}$ spatial masking. **d)** Enlarged version of c).

The same process of plotting the F1 score, MCC, balanced accuracy and visibility of a target from the OP7 data against the number of components, as in Figure 10, was applied to one of the target materials (green ceramic) present in the images from Figure 14. These graphs can be seen in Figure 15.

As in Figure 10 using the OP7 data, the average F1 score and MCC both increase with the number of retained components until $K = L$. ICA and FPCA both perform well on average at $K = 20$ whereas both PCA and MNF require more components to represent the data fully. Applying the spatial DR to each of the spectral DR methods improves both their F1 score and MCC regardless of the number of components retained. Similar to the results from Section 3.4, the balanced accuracy and visibility are lowered when using spatial DR because of the reduced TN class and more similar average background signature respectively. As in the results gathered from the OP7 dataset, applying spectral DR improves the balanced accuracy and visibility of the full spatial scene at lower values of K. The remaining results were obtained with $K = 20$ as it provided a good balance between detector performance
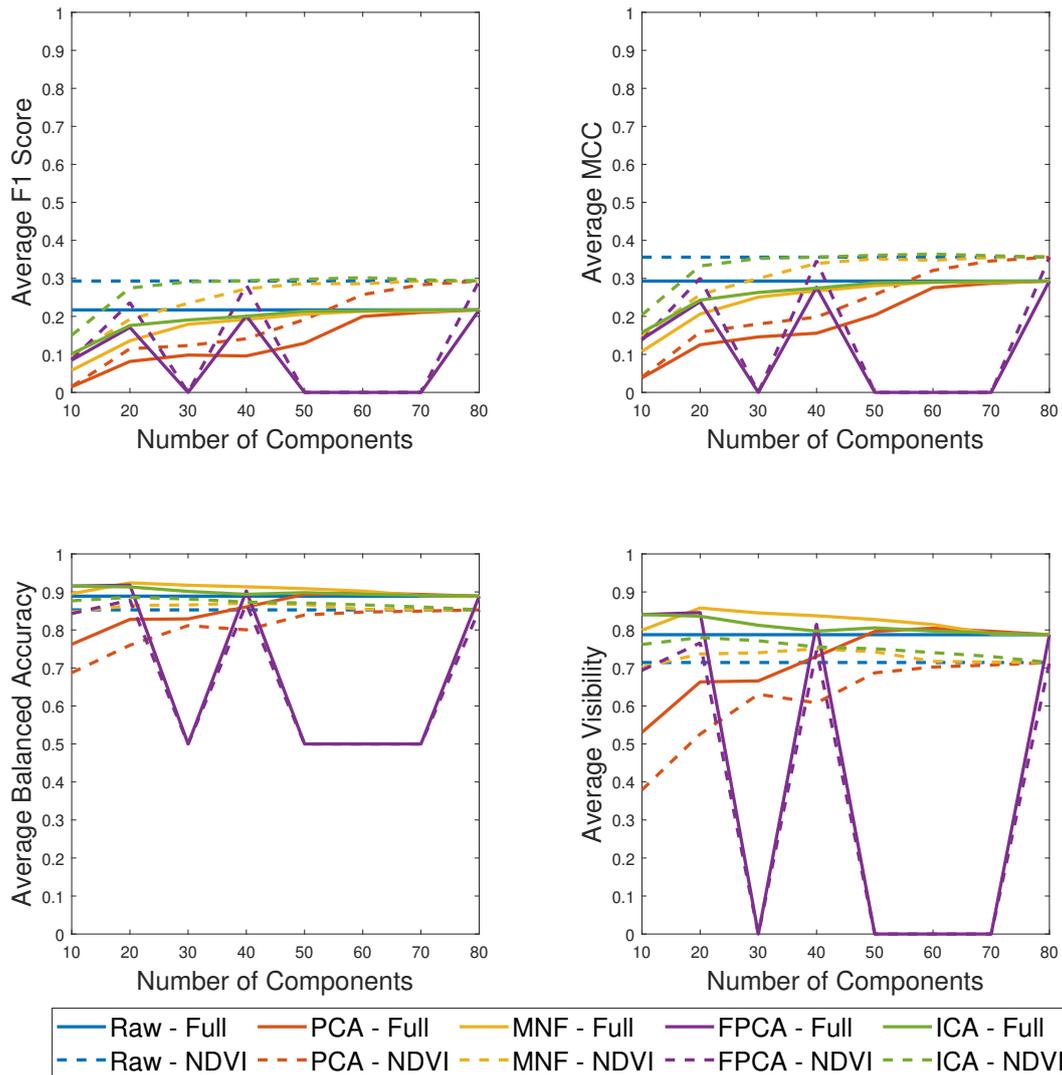
Figure 15. Performance measures using the ACE TD algorithm and each DR scheme with a varying number of retained components on the UDRC Selene dataset.

and compression. The results shown in Figure 15 also indicate that improved performance could be obtained at $K = 40$ at the expense of compression rate. It must again be noted that FPCA requires more careful consideration when selecting the value of K in order to avoid the creation of a singular matrix and avoid the dips in performance as seen in Figure 15. This is discussed in detail in Section 4.

Similar to the results obtained on the OP7 dataset in Figure 11, removing the vegetation and simplifying the background class improves separation between the synthetic background and targets. Whilst there is an overall decrease in target visibility, as the average spectra is more similar to the desired targets, there is less varied information to be represented, either in the full dimensionality image or in a dimensionality reduced subspace. This leads to less confusion in the detection image, as shown in Figure 16d, where the clutter present in the scene is less likely to be misidentified as a target, when compared to Figure 16b.
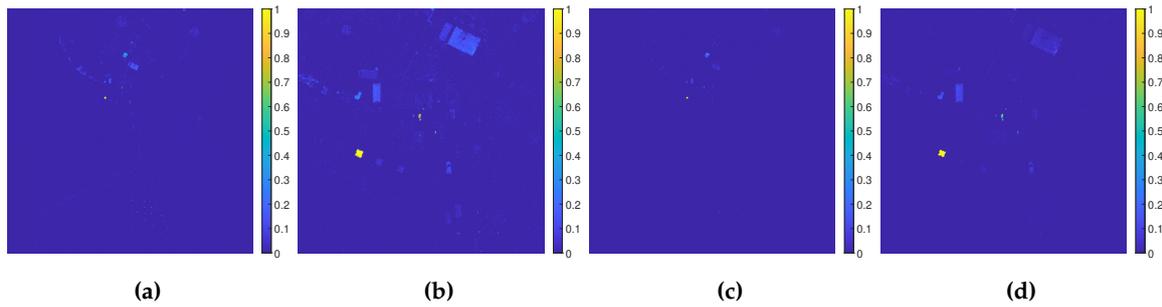
Figure 16. Detection statistics of the ACE algorithm on the full dimensionality data where yellow indicates a response of 1 and blue represents a response of 0. **a)** Without spatial DR. **b)** Enlarged version of a). **c)** With NDVI$_{re}$. **d)** Enlarged version of c).
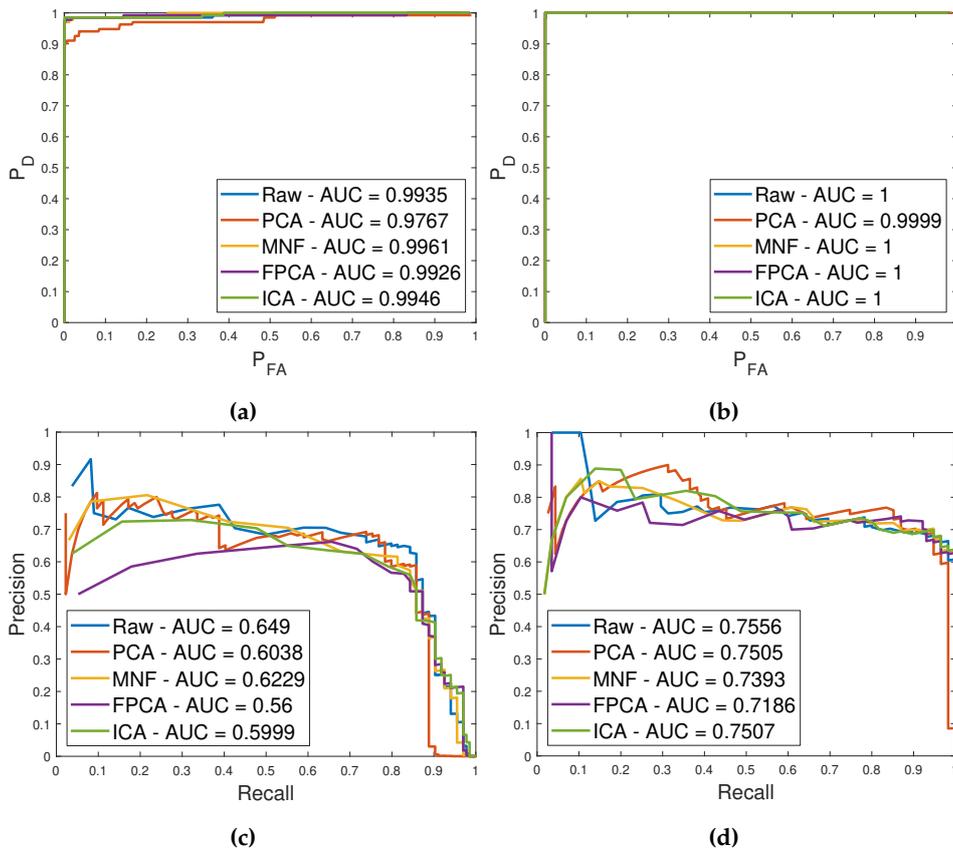
Figure 17. ROC and PR curves for detecting Target 3 (green ceramic) in the Selene dataset **a)** ROC curve using the full spatial dimensionality data. **b)** ROC curve using spatial DR pre-processing. **c)** PR curve using the full spatial dimensionality data. **d)** PR curve using spatial DR pre-processing.

The ROC curves in Figures 17a and 17b are of the ACE detection statistics on the full scene and NDVI$_{re}$-based spatial DR scene respectively. The two sets of ROC curves are almost identical and do not provide definitive results, but indicate a small improvement when using the spatial DR. Comparing the PR curves in Figures 17c and 17d show that when each spectral DR scheme is used in conjunction with spatial DR the AUC is increased by 10-15%.

The average results for each target in the Selene dataset are shown in Table 7 when considering the full scene and in Table 8 when spatial DR has been applied. The average performance of the ACE detector when combined with each spatial and spectral DR method used are shown in Figure 18.

Table 7. Average performance measures for each target in the Selene dataset without spatial DR applied using the ACE algorithm.

K = 20	DR	ACE-Full							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
Brown Carpet	Raw	**0.97**	0.33	0.65	0.19	0.67	0.82	0.17	0.24
	PCA	**0.97**	0.06	0.60	0.04	0.64	0.80	0.03	0.07
	MNF	**0.97**	0.46	**0.75**	0.17	**0.78**	**0.87**	0.15	0.21
	FPCA	**0.97**	0.55	**0.75**	0.22	**0.78**	**0.87**	0.20	0.26
	ICA	**0.97**	**0.57**	**0.75**	**0.25**	0.78	**0.87**	**0.22**	**0.28**
Green Carpet	Raw	0.98	**0.61**	0.82	**0.32**	0.83	0.91	**0.36**	**0.43**
	PCA	0.95	0.07	0.45	0.06	0.51	0.72	0.03	0.06
	MNF	**0.98**	0.54	0.80	0.23	0.83	0.89	0.24	0.30
	FPCA	0.98	0.58	0.85	0.25	0.89	**0.92**	0.29	0.35
	ICA	0.98	0.60	**0.86**	0.22	**0.90**	**0.92**	0.27	0.32
Green Ceramic	Raw	**0.99**	**0.65**	**0.94**	**0.19**	0.94	**0.96**	**0.29**	**0.39**
	PCA	0.98	0.60	0.85	0.16	0.89	0.92	0.19	0.26
	MNF	**0.99**	0.60	0.93	0.13	0.95	**0.96**	0.20	0.29
	FPCA	**0.99**	0.54	**0.94**	0.12	**0.96**	**0.96**	0.20	0.30
	ICA	**0.99**	0.52	**0.94**	0.13	**0.96**	**0.96**	0.20	0.30
Green Perspex	Raw	**1.00**	**0.63**	0.95	**0.22**	0.95	**0.97**	**0.32**	**0.42**
	PCA	0.99	0.44	0.91	0.08	0.93	0.95	0.12	0.20
	MNF	**1.00**	0.55	0.95	0.16	**0.97**	**0.97**	0.24	0.33
	FPCA	**1.00**	0.51	0.95	0.16	**0.97**	**0.97**	0.25	0.34
	ICA	**1.00**	0.57	**0.96**	0.15	**0.97**	**0.97**	0.23	0.33
Grey Ceramic	Raw	**0.99**	**0.61**	0.77	**0.31**	0.78	0.88	**0.27**	**0.34**
	PCA	0.98	0.47	0.81	0.13	0.83	0.90	0.11	0.17
	MNF	**0.99**	0.58	**0.85**	0.16	**0.88**	**0.92**	0.18	0.24
	FPCA	**0.99**	0.55	0.84	0.18	0.87	0.91	0.21	0.27
	ICA	**0.99**	0.53	0.82	0.18	0.85	0.90	0.19	0.25
Orange Perspex	Raw	**0.99**	**0.32**	0.90	**0.12**	0.90	0.95	**0.20**	**0.31**
	PCA	**0.99**	0.25	0.92	0.05	0.93	**0.96**	0.08	0.18
	MNF	**0.99**	0.29	**0.93**	0.07	**0.94**	**0.96**	0.13	0.24
	FPCA	**0.99**	0.30	**0.93**	0.08	**0.94**	**0.96**	0.14	0.25
	ICA	**0.99**	0.31	**0.93**	0.08	**0.94**	**0.96**	0.14	0.25
White Perspex	Raw	0.98	0.07	0.48	**0.07**	0.49	0.74	**0.05**	0.10
	PCA	**0.99**	**0.27**	**0.83**	0.04	**0.85**	**0.91**	**0.05**	**0.11**
	MNF	**0.99**	0.10	0.65	0.04	0.67	0.82	0.03	0.08
	FPCA	0.98	0.03	0.56	0.04	0.59	0.78	0.03	0.07
	ICA	0.98	0.02	0.45	0.03	0.49	0.72	0.02	0.05
All Spectra	Raw	0.99	**0.46**	0.77	**0.21**	0.78	0.88	**0.23**	**0.31**
	PCA	0.98	0.30	0.74	0.08	0.77	0.87	0.08	0.14
	MNF	**0.99**	**0.46**	**0.82**	0.14	**0.85**	**0.91**	0.16	0.23
	FPCA	**0.99**	0.45	**0.82**	0.16	0.84	0.90	0.19	0.26
	ICA	**0.99**	0.45	0.79	0.15	0.82	0.89	0.18	0.25

Table 8. Average performance measures for each target in the Selene dataset with NDVI$_{re}$-based spatial DR applied using the ACE algorithm.

| K = 20	DR	ACE-NDVI$_{re}$							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
Brown Carpet	Raw	**1.00**	0.20	0.53	0.15	0.55	0.76	0.13	0.18
	PCA	0.97	0.01	0.31	0.00	0.37	0.65	0.01	0.03
	MNF	0.99	0.15	0.56	0.12	0.61	0.78	0.09	0.14
	FPCA	0.99	0.27	0.59	0.17	0.64	0.79	0.13	0.18
	ICA	**1.00**	**0.47**	**0.68**	**0.22**	**0.73**	**0.84**	**0.19**	**0.25**
Green Carpet	Raw	**1.00**	**0.63**	0.82	**0.37**	0.83	0.91	**0.41**	**0.47**
	PCA	0.95	0.05	0.41	0.05	0.46	0.70	0.04	0.07
	MNF	**1.00**	0.48	0.75	0.24	0.79	0.87	0.24	0.31
	FPCA	**1.00**	0.58	0.82	0.30	0.86	0.91	0.32	0.38
	ICA	**1.00**	0.61	**0.89**	0.28	**0.93**	**0.94**	0.34	0.40
Green Ceramic	Raw	**1.00**	**0.70**	0.96	**0.37**	0.96	0.97	**0.49**	**0.56**
	PCA	**1.00**	0.63	0.91	0.25	0.93	0.95	0.30	0.38
	MNF	**1.00**	0.62	0.96	0.27	0.97	0.97	0.38	0.46
	FPCA	**1.00**	0.60	0.96	0.28	0.97	0.97	0.39	0.46
	ICA	**1.00**	0.62	**0.96**	0.30	**0.98**	**0.98**	0.42	0.49
Green Perspex	Raw	**1.00**	**0.68**	**0.97**	**0.41**	0.97	**0.98**	**0.54**	**0.60**
	PCA	**1.00**	0.61	0.96	0.23	0.98	0.98	0.31	0.38
	MNF	**1.00**	0.60	0.97	0.32	0.98	0.98	0.43	0.50
	FPCA	**1.00**	0.60	0.97	0.33	0.98	0.98	0.44	0.50
	ICA	**1.00**	0.64	0.97	0.33	**0.98**	0.98	0.44	0.51
Grey Ceramic	Raw	**1.00**	**0.62**	0.73	**0.40**	0.75	0.86	**0.35**	**0.41**
	PCA	**1.00**	0.47	0.76	0.22	0.79	0.87	0.19	0.26
	MNF	**1.00**	0.60	0.78	0.29	0.82	**0.89**	0.28	0.35
	FPCA	**1.00**	0.58	0.78	0.29	**0.83**	**0.89**	0.30	0.35
	ICA	**1.00**	0.60	**0.79**	0.28	**0.83**	**0.89**	0.27	0.34
Orange Perspex	Raw	**1.00**	0.35	0.87	**0.18**	0.87	0.93	**0.25**	**0.36**
	PCA	**1.00**	0.25	0.91	0.09	0.91	0.95	0.14	0.25
	MNF	**1.00**	0.35	0.92	0.11	0.93	0.95	0.18	0.29
	FPCA	**1.00**	0.33	0.92	0.12	0.92	0.95	0.18	0.30
	ICA	**1.00**	**0.36**	**0.93**	0.11	**0.94**	**0.96**	0.17	0.29
White Perspex	Raw	0.98	0.11	0.41	0.10	0.43	0.70	0.08	0.12
	PCA	**0.99**	**0.24**	**0.72**	**0.10**	**0.75**	**0.85**	**0.09**	**0.16**
	MNF	0.98	0.06	0.50	0.06	0.54	0.75	0.04	0.09
	FPCA	0.98	0.06	0.48	0.05	0.53	0.74	0.04	0.09
	ICA	0.94	0.02	0.35	0.02	0.41	0.67	0.02	0.05
All Spectra	Raw	**1.00**	0.46	0.73	**0.28**	0.74	0.86	**0.30**	**0.37**
	PCA	0.98	0.30	0.67	0.13	0.70	0.83	0.14	0.20
	MNF	0.99	0.40	0.75	0.20	0.78	0.87	0.22	0.28
	FPCA	0.99	0.43	0.76	0.22	0.79	**0.88**	0.24	0.31
	ICA	0.99	**0.47**	**0.77**	0.22	**0.80**	**0.88**	0.26	0.32
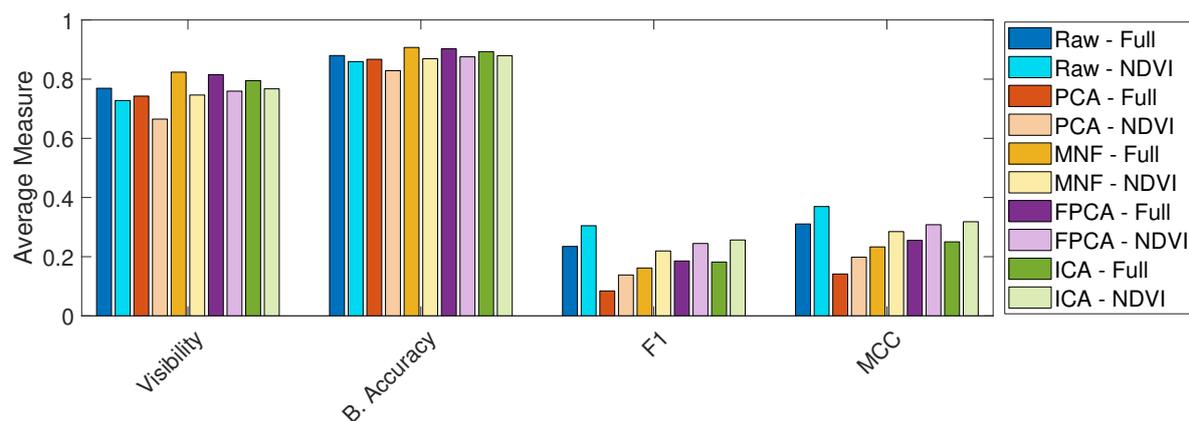
Figure 18. Comparison between the performance measures when combining spatial and spectral DR on the Selene dataset

In general, from Tables 7 and 8, the AUC of both the ROC and PR curves is similar regardless of the spectral and spatial DR used. By applying $NDVI_{re}$-based spatial DR, the precision of the spatial DR coupled methods increases in comparison to using the full spatial scene with the recall dropping slightly. As seen in Figure 15, there is a decrease in visibility of the target when using the spatial DR as the highly dissimilar vegetative background is removed. Figure 18 shows that, on average, the visibility drops slightly for each of the spectral DR methods when $NDVI_{re}$ is applied when $K = 20$ components are retained. The balanced accuracy also decreases slightly due to the reduction in the size of the TN class. Both the F1 score and MCC are improved when using spatial DR in all methods tested. The full dimensionality images with and without spatial DR have the best performance. However, of the spectral DR methods used, MNF, FPCA and ICA perform similarly despite retaining the equivalent of only 25% of the total spectral components. When combined with spatial DR both ICA and FPCA maintain the same level of performance as the full dimensionality image with no spatial DR applied. Applying the proposed method to the Selene dataset (Figure 18 and Tables 7 and 8) allows for improved results, however these improvements are not as significant as those achieved from the processing of the OP7 dataset. This can be attributed to the increased complexity of the Selene trial images when compared to the OP7 data. The performance can be improved further by retaining additional DR components, as shown in Figure 15, albeit at the expense of compression and therefore at an increased computational cost.

4. Discussion

The proposed $NDVI_{re}$-based spatial DR is relatively simple, requiring information from only two wavelengths and can be readily applied to TD and other similar applications. By using $NDVI_{re}$ it is possible to detect varied spectral targets composed of metals, plastics and other synthetic materials against a vegetative background. NDVI variants allow for the discrimination between vegetative and non-vegetative pixels due known material characteristics in the red-edge portion of the spectrum. Other VIs, whilst not considered here, as exploiting the red-edge portion was determined the key component of this method, may provide alternative insights and allow for the more optimal detection of additional materials in alternative environments. By combining both spatial and spectral DR, the computational complexity and memory requirements can be reduced whilst maintaining, or in some cases improving upon, detection performance as shown in Figures 13 and 18. Using spatial DR had little effect on the AUC of the ROC or PR curves, the main improvements came from the increased F1 score, MCC and precision. On average, there is a slight reduction in recall and balanced accuracy, however, one correctly detected and classified pixel per target may be sufficient for certain applications.

The complexity and performance of the spectral DR methods utilised varies. PCA is the simplest method used but also requires the most spectral components to be retained in order to be competitive. Applying the spatial DR and simplifying the background prior to performing spectral DR improved the

performance of all methods but most notably when using PCA, which was competitive in both datasets with the addition of spatial DR. MNF can be conceptualised as two PCAs, one for noise reduction and the second to transform the noise whitened data into the reduced subspace. This extra noise removal step offers a distinct advantage when compared to PCA and allows it to perform similarly to FPCA and ICA. FPCA performed favourably in both datasets and is efficient given the simplification when calculating the partial covariance matrix. However when using FPCA the choices of the number of components, K, and the height, H, and Width, W, of the folded matrix are far more sensitive than the other methods and are subject to two rules:

1. K must be a factor of the total number of wavelengths L

   or
2. When selecting the folding parameters H and W, $L > (H-1)W$

In any case where the first rule is true, the expression in the second rule will automatically be valid. H was selected to be half the value of K in order to adapt with the changing number of components. However, due to the folding process, padding the folded array with zeros in order to fulfil the expression $H \times W = L$, if these zeros formed an entire row of the covariance matrix they will form a zero component in both the projected image and target. When these interact in each of the TD algorithms, usually by inner product, it forms a singular matrix. As inverse matrices are prevalent in the implementations of the TD algorithms used, singular matrices completely suppresses the detection. This phenomenon caused the undulating behaviours present in Figures 10 and 15 and informed the choice of the number of DR components in order to compare each TD algorithm. ICA is the most complicated and computationally expensive method to compute, but performed well on both datasets. Only using the full dimensionality data, with and without spatial DR was an improvement over the ICA based methods. In general the spectral DR methods, whilst increasing the balanced accuracy and visibility when smaller numbers of components are retained, decrease the F1 score and MCC when compared with the raw full dimensionality data. Both FPCA and ICA offer consistent and improved detection when combined with ACE and NDVI$_{re}$-based spatial DR. In general, the most impressive results are obtained using the ACE TD algorithm which corroborates the conclusions of other similar works investigating this topic [11,14,17].

The methods detailed here offered improvement to the TD performance on both datasets considered. However greater improvements were obtained on the simpler dataset. Increasing the number of spectral DR components retained to account for the increased variability in the Selene dataset would improve the performance. This is at the expense of the compression rates that can be achieved at lower values of K. On average applying NDVI$_{re}$-based spatial DR increases precision and slightly decreases the recall of the TD algorithms used. The visibilities of the targets decrease as background pixels which are dissimilar to the targets are not considered. The average background signature, after applying the NDVI$_{re}$-based spacial DR, becomes more similar to the target signatures. However, applying spectral DR and mapping the data into a more informative subspace can alleviate this issue.

5. Conclusion

DR is a tool often employed in various hyperspectral imaging applications, usually to reduce the number of spectral bands present in an image due to its high spectral redundancy. However, known spatial redundancies are rarely exploited. This paper provides an investigation into how spatial DR can be utilised in a TD application. We have shown that in each case tested using multiple spectral DR schemes, the addition of a spatial DR pre-processing stage improved the performance of the TD algorithm considered. By applying both spatial and spectral DR the complexity of the data is reduced and computational cost and memory requirements can be lowered.

We used robust, classical TD/AD and DR algorithms in order to assess the proposed method. The provision of *a priori* information gives the TD algorithms an advantage over AD algorithms like the RXD

for example. Whilst the RXD correctly identifies the anomalous pixels, it fails to discriminate between specific target spectra resulting in low precision. Therefore, AD is insufficient for the application we are proposing. Of the detection methods tested, the ACE algorithm performs the best both when considering the full spatial scene and when applying the NDVI$_{re}$-based spatial DR - especially when combined with the FPCA and ICA DR algorithms.

We have shown that the proposed pipeline can compress an input image by >90% whilst maintaining the detector performance seen in the processing of the raw images. This pipeline is readily applicable in TD scenarios where the predominant background is comprised of vegetative pixels. The proposed method may be adapted to suppress other, highly predictable background signatures given an appropriate index. Indices such as the built-up index could provide the inverse to NDVI and its variants masking non-vegetative pixels directly, or alternatively providing auxiliary features. Additionally, multiple indices can be generated rapidly and combined to provide additional information about the pixels in a scene. Existing indices could also be used in the detection of camouflaged objects or bespoke alternative measures may be developed from such an investigation.

Potential future work includes using an adaptive method for selecting the optimal number of components, *K*, to retain in each DR method. In PCA, MNF and FPCA, variations on scree plots [20] can be used to find the elbow point. Alternatively, the value of *K* at which the number of components represent a sufficient percentage of the variance in the data could be chosen. Similarly for ICA, VD [28] can be used to estimate the number of spectrally distinct sources in the image and allows for the automation of this approach.

Although the proposed spatial DR approach has been tested on classical DR and TD/AD algorithms more state-of-the-art approaches to spectral DR could be considered as well as more complex detection algorithms. While the visibility of the target generally dropped when using spectral DR, the detection was improved and so a measure which can determine how distinctive the target is within the reduced subspace would be of benefit. Along with spectral DR other methods of spatial DR could be considered.

In order to avoid saturation of tables and results, the most informative and interesting results were included here. The full set of results generated from this work will be available online at a later date.

Acronyms

ACE Adaptive Cosine Estimator.
AD Anomaly Detection.
AUC Area Under the Curve.
CEM Constrained Energy Minimisation.
DR Dimensionality Reduction.
EVD Eigenvalue Decomposition.
FAR False Alarm Rate.
FN False Negative.
FP False Positive.
FPCA Folded Principal Component Analysis.
FPR False Positive Rate.
GSD Ground Sample Distance.
IC Independent Components.
ICA Independent Component Analysis.
MCC Matthew's Correlation Coefficient.
MNF Maximum Noise Fraction.
NDVI Normalised Difference Vegetation Index.
NDVI$_{re}$ Normalised Difference Vegetation Index (red-edge).
NIPALS Non-linear Iterative Partial Least Squares.
NIR Near-InfraRed.

OSP Orthogonal Subspace Projection.
P$_d$ Probability of Detection.
P$_{fa}$ Probability of False Alarm.
PC Principal Components.
PCA Principal Component Analysis.
PR Precision-Recall.
RENDVI Red Edge Normalised Difference Vegetation Index.
ROC Receiver Operator Characteristic.
RXD Reed-Xiaoli Detector.
SAM Spectral Angle Mapper.
SID Spectral Information Divergence.
SNR Signal-to-Noise Ratio.
TD Target Detection.
TN True Negative.
TNR True Negative Rate.
TP True Positive.
TPR True Positive Rate.
VD Virtual Dimensionality.
VI Vegetation Indices.
VNIR Visible and Near-InfraRed.

References

1. Van Aardt, J.A.; McKeown, D.; Faulring, J.; Raqueño, N.; Casterline, M.; Renschler, C.; Eguchi, R.; Messinger, D.; Krzaczek, R.; Cavillia, S. Geospatial disaster response during the Haiti earthquake: A case study spanning airborne deployment, data collection, transfer, processing, and dissemination. *Photogrammetric engineering and remote sensing* **2011**, *77*, 943–952.

2. Ettehadi Osgouei, P.; Kaya, S.; Sertel, E.; Alganci, U. Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sensing* **2019**, *11*. doi:10.3390/rs11030345.

3. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors* **2017**, *2017*.

4. Theiler, J.; Kucer, M.; Ziemann, A. Experiments in anomalous change detection with the Viareggio 2013 trial dataset. Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXVI. International Society for Optics and Photonics, 2020, Vol. 11392, p. 1139211.

5. Ghaderpour, E.; Vujadinovic, T. Change Detection within Remotely Sensed Satellite Image Time Series via Spectral Analysis. *Remote Sensing* **2020**, *12*, 4001.

6. Kwan, C.; Chou, B.; Yang, J.; Tran, T. Compressive object tracking and classification using deep learning for infrared videos. Pattern Recognition and Tracking XXX. International Society for Optics and Photonics, 2019, Vol. 10995, p. 1099506.

7. Manolakis, D.; Truslow, E.; Pieper, M.; Cooley, T.; Brueggeman, M. Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms. *IEEE Signal Processing Magazine* **2013**, *31*, 24–33.

8. Chang, C.I., Hyperspectral target detection. In *Real-Time Progressive Hyperspectral Image Processing*; Springer, 2016; pp. 131–172.

9. Wang, Q.; Lin, Q.; Li, M.; Tian, Q. A new target detection algorithm: spectra sort encoding. *International Journal of Remote Sensing* **2009**, *30*, 2297–2307. doi:10.1080/01431160802549351.

10. Nasrabadi, N.M. Hyperspectral target detection: An overview of current and future challenges. *IEEE Signal Processing Magazine* **2014**, *31*, 34–44.

11. Manolakis, D.; Lockwood, R.; Cooley, T.; Jacobson, J. Is there a best hyperspectral detection algorithm? Algorithms and technologies for multispectral, hyperspectral, and ultraspectral imagery XV. International Society for Optics and Photonics, 2009, Vol. 7334, p. 733402.

12. Chang, C.I.; Cao, H.; Chen, S.; Shang, X.; Yu, C.; Song, M. Orthogonal Subspace Projection-Based Go-Decomposition Approach to Finding Low-Rank and Sparsity Matrices for Hyperspectral Anomaly Detection. *IEEE Transactions on Geoscience and Remote Sensing* **2020**.

13. Chang, C.I. *Hyperspectral imaging: techniques for spectral detection and classification*; Vol. 1, Springer Science & Business Media, 2003.

14. Jin, X.; Paswaters, S.; Cline, H. A comparative study of target detection algorithms for hyperspectral imagery. Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV. International Society for Optics and Photonics, 2009, Vol. 7334, p. 73341W.

15. Olson, C.; Nichols, J.; Michalowicz, J.; Bucholtz, F. *Improved outlier identification in hyperspectral imaging via nonlinear dimensionality reduction*; Vol. 7695, *SPIE Defense, Security, and Sensing*, SPIE, 2010.

16. Manolakis, D.; Pieper, M.; Truslow, E.; Cooley, T.; Brueggeman, M.; Lipson, S. The remarkable success of adaptive cosine estimator in hyperspectral target detection. Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX. International Society for Optics and Photonics, 2013, Vol. 8743, p. 874302.

17. Bakken, S.; Orlandic, M.; Johansen, T.A. The effect of dimensionality reduction on signature-based target detection for hyperspectral remote sensing. CubeSats and SmallSats for Remote Sensing III. International Society for Optics and Photonics, 2019, Vol. 11131, p. 111310L.

18. Macfarlane, F.; Murray, P.; Marshall, S.; White, H. A Fast Hyperspectral Hit-or-Miss Transform With Integrated Projection-Based Dimensionality Reduction. Hyperspectral Imaging HSI, 2018.

19. Macfarlane, F.; Murray, P.; Marshall, S.; White, H. *Object detection and classification in aerial hyperspectral imagery using a multivariate hit-or-miss transform*; Vol. 10986, *SPIE Defense + Commercial Sensing*, SPIE, 2019.

20. Abdi, H.; Williams, L.J. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics* **2010**, *2*, 433–459.

21. Andrecut, M. Parallel GPU implementation of iterative PCA algorithms. *Journal of Computational Biology* **2009**, *16*, 1593–1599.

22. Green, A.A.; Berman, M.; Switzer, P.; Craig, M.D. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on geoscience and remote sensing* **1988**, *26*, 65–74.

23. Luo, G.; Chen, G.; Tian, L.; Qin, K.; Qian, S.E. Minimum Noise Fraction versus Principal Component Analysis as a Preprocessing Step for Hyperspectral Imagery Denoising. *Canadian Journal of Remote Sensing* **2016**, *42*, 106–116. doi:10.1080/07038992.2016.1160772.

24. Zabalza, J.; Ren, J.; Yang, M.; Zhang, Y.; Wang, J.; Marshall, S.; Han, J. Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* **2014**, *93*, 112–122.

25. Cao, L.J.; Chua, K.S.; Chong, W.K.; Lee, H.P.; Gu, Q.M. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. *Neurocomputing* **2003**, *55*, 321–336. doi:https://doi.org/10.1016/S0925-2312(03)00433-8.

26. Hyvärinen, A.; Oja, E. A fast fixed-point algorithm for independent component analysis. *Neural computation* **1997**, *9*, 1483–1492.

27. Cardoso, J.F.; Souloumiac, A. Blind beamforming for non-Gaussian signals. IEE proceedings F (radar and signal processing). IET, 1993, Vol. 140, pp. 362–370.

28. Chang, C. A Review of Virtual Dimensionality for Hyperspectral Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2018**, *11*, 1285–1305. doi:10.1109/JSTARS.2017.2782706.

29. Wang, J.; Chang, C. Applications of Independent Component Analysis in Endmember Extraction and Abundance Quantification for Hyperspectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing* **2006**, *44*, 2601–2616. doi:10.1109/TGRS.2006.874135.

30. Rouse Jr, J.W.; Haas, R.H.; Schell, J.; Deering, D. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation **1973**.

31. Hansen, P.; Schjoerring, J. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote sensing of environment* **2003**, *86*, 542–553.

32. Gitelson, A.; Merzlyak, M.N. Spectral reflectance changes associated with autumn senescence of Aesculus hippocastanum L. and Acer platanoides L. leaves. Spectral features and relation to chlorophyll estimation. *Journal of plant physiology* **1994**, *143*, 286–292.

33. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote sensing of environment* **2002**, *81*, 337–354.

34. Kraut, S.; Scharf, L.L.; Butler, R.W. The adaptive coherence estimator: a uniformly most-powerful-invariant adaptive detection statistic. *IEEE Transactions on Signal Processing* **2005**, *53*, 427–438. doi:10.1109/TSP.2004.840823.

35. Kruse, F.A.; Lefkoff, A.; Boardman, J.; Heidebrecht, K.; Shapiro, A.; Barloon, P.; Goetz, A. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote sensing of environment* **1993**, *44*, 145–163.

36. Chang, C.I. Spectral information divergence for hyperspectral image analysis. IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No. 99CH36293). IEEE, 1999, Vol. 1, pp. 509–511.

37. Reed, I.S.; Yu, X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **1990**, *38*, 1760–1770.

38. Chang, C.I.; Chen, J. Orthogonal Subspace Projection Using Data Sphering and Low-Rank and Sparse Matrix Decomposition for Hyperspectral Target Detection. *IEEE Transactions on Geoscience and Remote Sensing* **2021**, pp. 1–19. doi:10.1109/TGRS.2021.3053201.

39. Chang, C.I. Orthogonal subspace projection (OSP) revisited: A comprehensive study and analysis. *IEEE transactions on geoscience and remote sensing* **2005**, *43*, 502–518.

40. Shen, S.S.; Bajorski, P.; Lewis, P.E. Impact of missing endmembers on the performance of the OSP detector for hyperspectral images, 2007. doi:10.1117/12.720101.

41. Bajorski, P. Analytical Comparison of the Matched Filter and Orthogonal Subspace Projection Detectors for Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing* **2007**, *45*, 2394–2402. doi:10.1109/TGRS.2007.896544.

42. Bajorski, P. Target Detection Under Misspecified Models in Hyperspectral Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2012**, *5*, 470–477. doi:10.1109/JSTARS.2012.2188095.

43. Gross, W.; Boehler, J.; Schilling, H.; Middelmann, W.; Weyermann, J.; Wellig, P.; Oechslin, R.; Kneubühler, M. Assessment of target detection limits in hyperspectral data. Target and Background Signatures. International Society for Optics and Photonics, 2015, Vol. 9653, p. 96530J.

44. Kerekes, J. Receiver operating characteristic curve confidence intervals and regions. *IEEE Geoscience and Remote Sensing Letters* **2008**, *5*, 251–255.

45. Saito, T.; Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS one* **2015**, *10*, e0118432.

46. Powers, D.M. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061* **2020**.

47. Matthews, B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1975**, *405*, 442–451. doi:https://doi.org/10.1016/0005-2795(75)90109-9.

48. Brodersen, K.H.; Ong, C.S.; Stephan, K.E.; Buhmann, J.M. The balanced accuracy and its posterior distribution. 2010 20th international conference on pattern recognition. IEEE, 2010, pp. 3121–3124.

49. Verdoja, F.; Grangetto, M. Graph Laplacian for image anomaly detection. *Machine Vision and Applications* **2020**, *31*, 1–16.