

doi: 10.6046/zrzyyg.2022093

引用格式: 张国建,刘胜震,孙英君,等. 基于弱监督鲁棒性自编码的高光谱异常检测[J]. 自然资源遥感,2023,35(2):167-175. (Zhang G J, Liu S Z, Sun Y J, et al. Hyperspectral anomaly detection based on the weakly supervised robust autoencoder[J]. Remote Sensing for Natural Resources, 2023, 35(2): 167-175.)

基于弱监督鲁棒性自编码的高光谱异常检测

张国建¹, 刘胜震², 孙英君¹, 俞凯杰³, 刘丽娜⁴

(1. 山东建筑大学测绘地理信息学院, 济南 250101; 2. 自然资源部第一大地测量队, 西安 710054; 3. 浙江中测新图地理信息技术有限公司, 湖州 313200; 4. 宁波市鄞州区测绘院, 宁波 315041)

摘要: 高光谱异常检测因其以无监督方式检测目标的能力而受到特别关注, 自动编码器及其变体可以自动提取深层特征, 还可以检测异常目标。由于训练集中存在异常, 自动编码器泛化性较强, 从而降低了从背景中区分异常的能力。为解决上述问题, 该文提出一种基于弱监督鲁棒性自编码的异常探测算法。首先提出了一种显著类别搜索策略, 采用基于概率的类别阈值来标记粗样本, 为网络弱监督学习做准备; 同时, 构建一个具有 $l_{2,1}$ 范数与异常-背景光谱距离共同约束的鲁棒性自编码网络框架, 该框架在训练期间对噪声和异常具有鲁棒性; 最后, 采用所有样本得到的重构误差检测异常目标。在4个高光谱数据集上进行实验, 结果表明, 与其他先进的异常检测算法相比, 所提算法具有更好的检测性能。

关键词: 高光谱图像; 异常检测; 显著类别搜索; 鲁棒性 AE

中图法分类号: TP 79 **文献标志码:** A **文章编号:** 2097-034X(2023)02-0167-09

0 引言

成像光谱仪的出现与发展促进了遥感技术的进步, 使得高光谱遥感技术在遥感领域受到越来越多的关注。高光谱影像能够同时获取地球表面空间信息和丰富的光谱信息, 这些地物光谱信息在目标探测与地物分类中具有显著优势^[1]。目标探测技术是高光谱遥感领域的一个重要应用, 而异常探测技术由于不需要任何目标与背景的先验信息, 因此实际应用价值更高, 广泛应用于精准农业、环境监测、军事领域等^[2]。异常目标定义为目标光谱特征与周围地物的光谱特征具有明显差异, 且其目标小、数量少、分布随机的特性, 因此异常目标的类型往往难以获取^[3]。高光谱异常检测通常具有以下特点^[4]: ①异常目标具有低概率特性; ②异常在光谱特征上与背景明显不同; ③没有关于背景或异常的光谱先验信息可用; ④在复杂背景情况下, 异常目标与背景混合, 往往为混合像元或亚像元。这些特性使得高光谱异常检测成为遥感领域中一个研究热点。

近年来, 研究者提出大量异常探测算法。由Reed等提出的RX算法^[5]通过假设背景服从某种

分布, 再使用统计方法来估计像素属于异常目标的概率。随后部分研究学者提出一些基于RX算法的改进算法, 例如非线性的核RX异常探测算法^[6]、基于权重RX探测算法^[7]等。然而, 真实高光谱影像含有多类背景地物且分布不均匀, 因此很难准确对不同背景地物建模, 在实际应用中往往受限^[8]。为了避免估计背景分布, Li等^[9]提出基于协同表示的异常探测算法(collaborative representation-based detector, CRD), 该算法假设每个像素可以由背景字典进行线性组合, 其中预测值与真值之间的残差被视为像素的异常程度。针对背景字典中存在异常点污染导致检测精度下降的问题, 刘万军等^[10]利用背景字典与字典均值的距离对背景字典中的各像元自适应设计权重, 通过降低异常像元的权重提升检测精度。基于低秩和稀疏矩阵分解的马氏距离异常检测算法(low-rank and sparse matrix decomposition-based Mahalanobis distance anomaly detection, LS-MAD)通过在去除稀疏异常后充分探索背景的低秩特性来建立背景模型^[11]。基于分数傅里叶熵的检测器(FrFE)创新性地将分数傅里叶熵理论引入高光谱异常检测中, 取得了良好的检测结果^[12]。尽管上述方法取得了良好的探测效果, 但仅利用了低阶

收稿日期: 2022-03-21; 修订日期: 2022-08-16

第一作者: 张国建(1989-), 男, 讲师, 主要从事摄影测量与遥感技术方面的研究。Email: 494088845@qq.com。

通信作者: 刘胜震(1987-), 男, 工程师, 主要从事测绘方面的研究。Email: szliu1987@126.com。

特征,这也限制了检测器的性能。

近年来,深度学习已经应用在高光谱遥感影像分类及目标探测当中,其中自编码网络(autoencoder, AE)是非监督学习领域中的一种,其以自动的方式获取抽象的、层次化的特征,已应用到高光谱异常检测中^[13-15]。Xie 等^[16]提出了一种光谱约束对抗 AE(SC_AAE),其中将光谱约束策略结合到对抗性 AE 中以学习高光谱数据的潜在表示,然后使用双层架构实现异常检测;Lu 等^[17]提出了基于流形约束 AE 网络,该方法添加嵌入流形学习约束,最后,由全局和局部组成的重构误差用于异常检测。为了提出异常检测性能,Zhao 等^[18]将高光谱影像分解为低秩部分和稀疏部分,低维潜在特征分别由 2 个堆叠 AE 学习,最后采用局部马氏距离检测异常目标。上述深度方法有效提升了检测性能,AE 网络在经过适当训练的情况下表现良好,而且由于没有背景或异常的先验光谱特征,训练集容易被异常像元破坏。当训练集中存在异常时,AE 学习性能非常好,以至于它可以很好地学习异常并以较小的重建误差重建异常,使得检测性能降低。

为解决上述问题,本文提出一种基于弱监督鲁棒性自编码的异常探测算法(weakly supervised ro-

bust AE,WSRAE)。该算法首先提出了一种显著类别搜索方法来搜索样本的伪标签,即弱监督学习所需的不准确标签。同时,引入 $l_{2,1}$ 范数与异常-背景光谱距离构建鲁棒性 AE(robust AE, RAE) 框架,以减轻异常部分对参数微调的影响,这使得 AE 网络在训练过程中更加对噪声和异常更具有鲁棒性。最后,将所有光谱样本输入到 RAE 模型中,从而产生具重构高光谱影像,利用重构误差以探测异常目标。实验证明,本文方法有效抑制了异常像元对背景建模干扰,同时降低了虚警率,提升了检测精度。

1 方法原理

针对训练样本中含有异常目标且 AE 网络训练不稳定问题,本文提出了一种用于高光谱异常检测的弱监督方法,并通过以弱监督方式对所提出的 RAE 网络进行对抗训练,从而学习具有背景鲁棒性和异常突出性的端到端重建。图 1 表示本文提出的 WSRAE 算法的框架,该方法由 3 个主要部分组成:显著类别搜索、弱监督 RAE 学习和重构影像的异常检测。

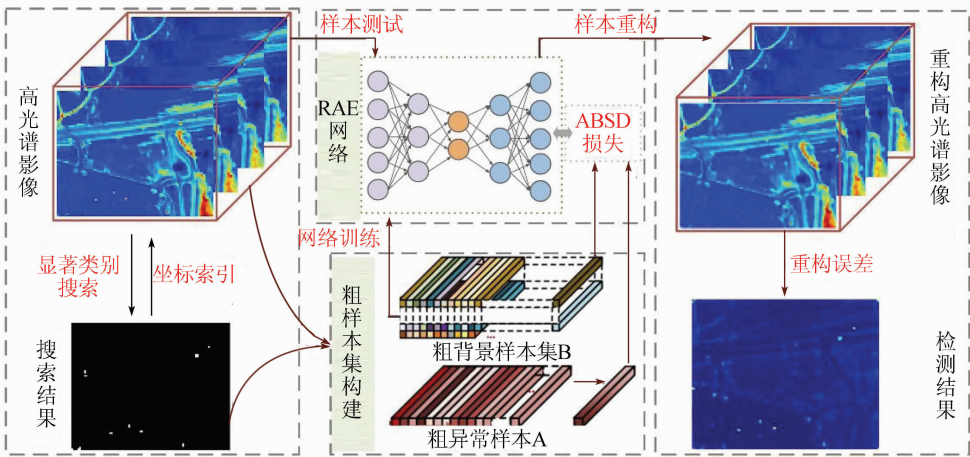


图 1 WSRAE 算法结构
Fig. 1 WSRAE algorithm structure

1.1 显著类别搜索

针对训练样本中含有异常目标导致异常检测精度降低,结合基于密度的噪声应用空间聚类(density-based spatial clustering of applications with noise, DBSCAN)算法^[19],本文设计一种非监督的背景与异常目标的显著类别搜索算法,其是为弱监督网络学习准备,旨在预测给定输入的伪标签。根据预测的伪标签,该过程搜索包含 2 个子集,包括一个粗略的异常样本集,其中像素具有较高的属于异常的概率,以及一个粗略的背景样本集,其中像素具有较高的属

于背景的概率。因此,显著类别搜索问题可以看作是一个初始异常-背景分离问题,该算法分为 3 个主要部分:显著类别提取、坐标索引和粗样本集构建。

在没有任何背景与目标先验知识的情况下,本文首先应用 DBSCAN 的非监督聚类方法获取类别概率图(即聚类结果图)。由于 DBSCAN 算法能够发现具有不同大小和形状的簇,并且在噪声和异常值的情况下对指定簇数的聚类性能更具有鲁棒性,因此 DBSCAN 在高光谱异常检测中具有显著优势。

在聚类过程中,利用 DBSCAN 扫描每个像元在 Eps (即相邻点之间的最小距离)距离内的邻域,若中心像素的邻居数超过 MinPts (最小点数),就形成一个簇,其中距离判定方法采用欧氏距离。

给定一个输入高光谱影像 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M \times N}]$, 其中 $\mathbf{y}_i \in \mathbb{R}^{L \times 1}$, L 为波段数, M 与 N 分别为行列数。在 (Eps, MinPts) 的条件下,得到类别概率图 $\mathbf{P} = \{p_i\}_{i=1}^{M \times N} \in \mathbb{R}^{M \times N \times 1}$ 。针对不同高光谱影像的聚类类别数量差异很大问题,结合背景样本出现在异常检测中的概率,本文构建一种基于概率的类别阈值方法,其目的是获得粗略的标签划分,而不是精细分类。首先通过设定阈值 ε , 大于阈值 ε 属于背景的样本的类别标签为 0, 而其他样本可以粗略地分类为类别标签为 1 的异常,表达式为:

$$\mathbf{S} = \begin{cases} s_i = 0 & (p_i \geq \varepsilon) \\ s_i = 1 & (p_i < \varepsilon) \end{cases}, \quad (1)$$

式中: $\mathbf{S} = \{s_i \mid i = (i_m, i_n)\}_{i_m=1, i_n=1}^{i_m=M, i_n=N}$, 表示类别显著图,作为 $\mathbf{Y}$ 的伪类别标签集; (i_m, i_n) 为第 i 个像素点在二维图像空间中的位置坐标,即 $i = N \cdot (i_m - 1) + i_n$ 。根据类别显著图 $\mathbf{S}$ 的坐标索引,搜索粗异常样本集 $\mathbf{A}$ 和粗背景样本集 $\mathbf{B}$ 为:

$$\mathbf{A} = \{\mathbf{y}_i \mid s_{i=(i_m, i_n)} = 1\}, \quad (2)$$

$$\mathbf{B} = \{\mathbf{y}_i \mid s_{i=(i_m, i_n)} = 0\}, \quad (3)$$

式中: 粗异常样本集 $\mathbf{A} = \{\mathbf{a}_i \in \mathbb{R}^{L \times 1}\}_{i=1}^{i=C_A}$; 粗背景样本集 $\mathbf{B} = \{\mathbf{b}_i \in \mathbb{R}^{L \times 1}\}_{i=1}^{i=C_B}$; C_A 与 C_B 分别表示样本集 $\mathbf{A}$ 与 $\mathbf{B}$ 的样本数,且 $C_A + C_B = M \times N$ 。对于粗略异常样本或背景样本集,可以容忍一些类别的判别错误,即 $\mathbf{S}$ 中预测的粗略标签并不总是真实的。

$$L_r = \|\widehat{\mathbf{B}} - \mathbf{B}\|_{2,1} = \sum_{i=1}^{C_B} \|\widehat{\mathbf{b}}_i - \mathbf{b}_i\|_2 = \text{Tr}[(\widehat{\mathbf{B}} - \mathbf{B})^T \mathbf{P}(\widehat{\mathbf{B}} - \mathbf{B})], \quad (7)$$

式中: $\mathbf{P}$ 为一个对角矩阵 $\mathbf{P}_{ii} = (1/\|\widehat{\mathbf{b}}_i - \mathbf{b}_i\|_2 + \varepsilon)$; ε 为一个很小的常数项。由于 RAE 在反向传播之前, $l_{2,1}$ 范数对每个样本的重建误差进行了归一化,因此网络能够学习更多关于背景的信息并抑制异常对网络微调的影响。所以 $l_{2,1}$ 范数约束的 RAE 网络可以更好地处理训练过程中的少部分异常样本。

此外,为了更好地区别异常和背景之间的特征,本文构建异常-背景光谱距离(anomaly-background spectral distance, ABSD)约束,其通过以最大化粗异常样本与粗背景样本之间的距离。ABSD 的值越大,2 个光谱向量的区分度越高。因此,集合 $\mathbf{A}$

因此,将粗背景样本集送入网络模型训练,即以弱监督方式训练,从而提升异常检测的性能。

1.2 RAE 网络

AE 网络是一种非监督学习的网络模型,其由编码器与解码器 2 部分组成,前者用来将输入数据转换为潜在特征,后者是将潜在特征生成为与原始输入数据相似的重构数据^[20]。在 AE 中,编码器的目标是将输入层 $\mathbf{x}$ 经过权重矩阵 $\mathbf{W}$ 和偏置 $\mathbf{b}$ 转换为隐藏层 $\mathbf{z}$, 公式为:

$$\mathbf{z} = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b}), \quad (4)$$

式中 σ 为激活函数。解码器利用权重矩阵 $\mathbf{W}'$ 和偏置 $\mathbf{b}'$ 将隐藏层 $\mathbf{z}$ 映射到重构层 $\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = \sigma(\mathbf{W}'\mathbf{z} + \mathbf{b}'), \quad (5)$$

AE 以神经网络作为函数逼近器,那么数据 $\mathbf{x}$ 的重构误差损失函数为:

$$L_r = \|\mathbf{x} - \hat{\mathbf{x}}\|_2. \quad (6)$$

由于 AE 网络在每个待测像素中都采用 l_2 范数的平方,因此噪声或异常像素往往会获得更大的重建误差,这将破坏 AE 网络的重构性能,因此无法判别一些异常像素。通常异常目标受到混合像元的影响,上述显著类别样本无法完整地将所有异常样本准确提取,保留了少部分异常样本位于粗略背景样本集中。 $l_{2,1}$ 范数在算法中更加有效应对噪声和异常值,并已广泛应用于许多领域^[21]。受此启发,本文首先引入 $l_{2,1}$ 范数作为目标函数,构建了一个新颖的 RAE 网络,同时以弱监督的方式来学习 RAE 网络,挖掘和增强异常与背景之间的判别特征。 L_r 目标函数定义如下:

和 $\mathbf{B}$ 之间的光谱距离 L_{ABSD} 目标函数定义如下:

$$L_{\text{ABSD}} = \|\mathbf{a}^T \mathbf{P}_{\mathbf{b}_i}^\perp \mathbf{a} + \mathbf{b}_i^T \mathbf{P}_{\mathbf{a}}^\perp \mathbf{b}_i\|_2, \quad (8)$$

式中: $\mathbf{P}_k^\perp = \mathbf{I} - k(k^T k)^{-1} k^T$, $\mathbf{P}_k^\perp$ 为 k 的正交子空间。此外,将粗异常样本集 $\mathbf{A}$ 压缩为异常样本向量 $\mathbf{a}$, 公式为:

$$\mathbf{a} = \frac{1}{C_A} \sum_{i=1}^{i=C_A} \mathbf{a}_i, \quad (9)$$

式中 $\mathbf{a} \in \mathbb{R}^{L \times 1}$ 为粗异常样本的平均值。

通过将公式 (7) 和 (8) 的 2 个目标函数与正的权衡参数 λ 集成在一起,RAE 网络的目标函数如下:

$$L = L_r + \lambda L_{\text{ABSD}} = \frac{1}{2M} \sum_{i=1}^{C_B} \|\widehat{\mathbf{b}}_i - \mathbf{b}_i\|_2 + \frac{\lambda}{2N} \sum_{i=1}^{M \times N} \|\mathbf{a}^T \mathbf{P}_{\mathbf{b}_i}^\perp \mathbf{a} + \mathbf{b}_i^T \mathbf{P}_{\mathbf{a}}^\perp \mathbf{b}_i\|_2, \quad (10)$$

式中：第一部分表示重构误差的 $l_{2,1}$ 范数，促使 RAE 网络学习更多的背景信息，并对噪声和异常目标更具鲁棒性；第二部分表示异常和背景之间的区别特征的 ABSD 光谱距离，以增强背景与异常的可分性。

1.3 WSRAE 算法

利用粗背景样本集对 RAE 网络训练完成后，利用得到训练完成的模型参数 $F_\theta(\cdot)$ 进行异常目标检测。将输入高光谱影像 $Y = [y_1, y_2, \cdots, y_{M \times N}]$ 输入至 RAE 网络，使用式(10)的 RAE 网络的模型参数 $F_\theta(\cdot)$ 获得重构的高光谱影像 $\hat{Y}$ ，公式为：

$$\hat{Y} = F_\theta(Y)。$$

(11)

通过上述重建过程，模型 $F_\theta(\cdot)$ 在从背景中识别异常时，使重建的 $\hat{Y}$ 比原始 Y 更具判别性。因此，最终的异常检测结果使用式(11) 计算，即重构误差，公式为：

$$D = \|Y - \hat{Y}\|_2。$$

(12)

因此，WSRAE 算法以弱监督的方式进行背景学习，其不仅可以生成分布均匀的背景，同时结合构建的 $l_{2,1}$ 范数损失与 ABSD 光谱距离损失，以增强异常与背景的分离。

- 本文所提的 WSRAE 算法的探测过程如下：
- 输入：高光谱影像 $Y = [y_1, y_2, \cdots, y_{M \times N}]$ ，阈值 ε ，权衡参数 λ 。
- 输出：高光谱异常检测图。
- 1) 显著类别搜索。
- ①利用 DBSCAN 算法，根据式(1)，进行显著类别提取；

- ②利用 $i = N \cdot (i_m - 1) + i_n$ ，进行坐标索引；
- ③根据式(2)—(3)，构建粗背景样本集 B 与粗异常样本集 A ；
- 2) RAE 模型训练。
- ①初始化，RAE 网络模型参数；
- ②输入粗样本集 B ，异常样本向量 a ；
- ③迭代训练，根据式(10)更新网络参数；
- 3) 异常检测。
- 根据式(11)—(12)，计算重构误差。

2 实验结果与分析

2.1 实验数据

在本文中，使用 4 个高光谱影像数据集来验证所提出算法的有效性，伪彩色图像和异常目标地物真值图，如图 2 所示。如图 2(a)—(b) 所示，Urban 数据集采用 HYDICE 传感器获取，空间分辨率为 1 m，影像尺寸为 80 像素 $\times$ 100 像素，具有 162 个光谱波段。如图 2(c)—(d) 所示，Gulfport 数据集采用 AVIRIS 传感器获取，空间分辨率为 3.4 m，影像尺寸为 100 像素 $\times$ 100 像素。在 400 ~ 2 500 nm 的波长范围内去除水汽波段与噪声波段后，具有 191 个波段。该数据集共有 60 个异常像素，异常地物为飞机。如图 2(e)—(f) 所示，San Diego - 1 数据集采用 AVIRIS 传感器获取，空间分辨率为 7.1 m，影像尺寸为 100 像素 $\times$ 100 像素，具有 189 个光谱波段。3 架飞机被视为异常。如图 2(g)—(h) 所示，San Diego - 2 数据集采用 AVIRIS 传感器获取，空间分辨率 20 m，影像尺寸为 100 像素 $\times$ 100 像素。

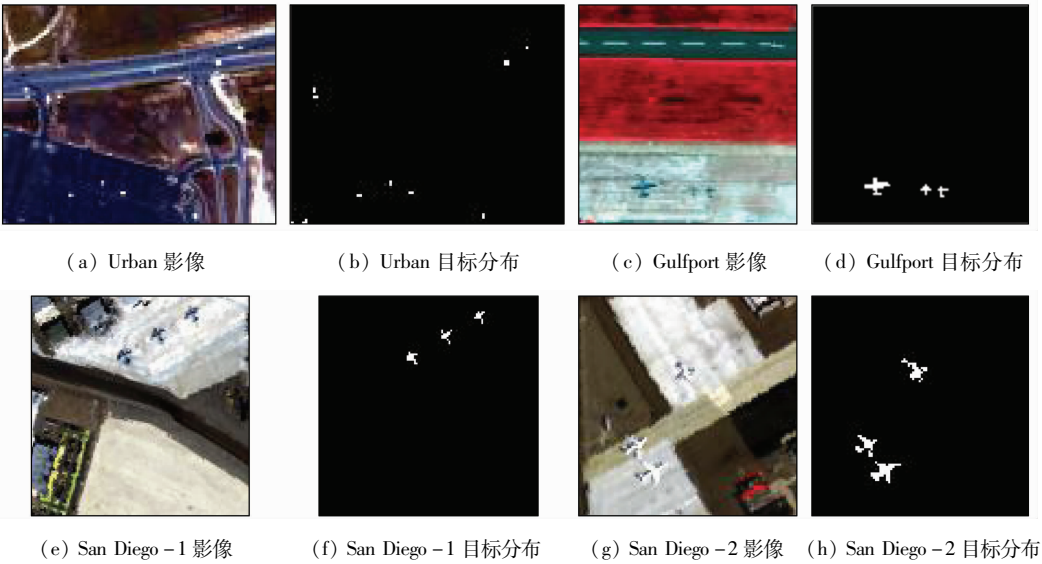


图 2 4 个高光谱数据集的伪彩色图像和对应的地物真值图
Fig.2 Pseudo color images and the corresponding ground truth maps of the 4 HSI datasets

去除与吸水区域相对应的不良波段后,包含 189 个波段。由 134 个像素组成的 3 架飞机被视为异常。

2.2 模型结构对比与分析

为验证本文提出的 WSRAE 算法的先进性,进行以下消融研究:①采用 l_2 范数约束的 AE 网络,即 AE(l_2);②采用 $l_{2,1}$ 范数约束的 AE 网络,即 AE($l_{2,1}$);③采用 l_2 范数与 ABSD 光谱距离共同约束的 AE 网络,即 AE(*ABSD*);④采用 $l_{2,1}$ 范数与 ABSD 光谱距离共同约束的 AE 网络,即本文提出的 RAE 网络。同时进行显著类别搜索构建粗背景样本集与粗异常样本集,对上述 4 种对比模型采用弱监督方式的方式进行训练。表 1 对比了各个模型的 AUC 的(P_d,P_f)和(P_f,τ)值,其中, P_d 为算法的检测率, P_f 为虚警率,通过设定不同的阈值 τ 计算对应的虚警率与检测率。AUC (P_d,P_f)值越接近

1,其检测性能越好,AUC (P_f,τ)值越接近 0,其误检的概率越低。从中可以看出,相比于传统采用 l_2 范数约束的 AE,采用 $l_{2,1}$ 范数的 AE 或者添加 ABSD 光谱距离约束的 AE 显著提升了异常检测精度,同时误检的概率也在降低。RAE 网络采用的 $l_{2,1}$ 范数与 ABSD 光谱距离共同约束的 AE,检测精度比原始 l_2 范数、 $l_{2,1}$ 范数以及 $l_{2,1}$ 范数与 ABSD 共同约束的 AE 检测精度大幅提升。从引入弱监督训练来看,相比非监督的网络检测结果,基于弱监督的网络检测结果在一定的程度上提升了检查精度,同时也降低了虚警率,这表明采用弱监督学习对高光谱异常检测有积极影响。总之,本文所提出的显著类别搜索方法的弱监督训练策略,以及引入 $l_{2,1}$ 范数与 ABSD 光谱距离构建 RAE 框架是有效的,表明了所设计方法的先进性和优越性。

表 1 4 组数据集中消融研究的 AUC 值
Tab.1 AUC values of ablation studies in four data sets

消融方案	Urban		Gulfport		San Diego - 1		San Diego - 2	
	(P_d,P_f)	(P_f,τ)	(P_d,P_f)	(P_f,τ)	(P_d,P_f)	(P_f,τ)	(P_d,P_f)	(P_f,τ)
AE(l_2)	0.977 0	0.034 2	0.969 0	0.020 3	0.959 8	0.027 3	0.946 7	0.038 2
WSAE(l_2)	0.980 5	0.025 5	0.974 5	0.032 5	0.980 5	0.019 6	0.979 0	0.013 4
AE($l_{2,1}$)	0.977 5	0.024 8	0.984 2	0.018 5	0.979 7	0.014 5	0.989 5	0.025 1
WSAE($l_{2,1}$)	0.985 4	0.019 2	0.982 8	0.013 5	0.982 5	0.025 2	0.974 0	0.015 7
AE(<i>ABSD</i>)	0.984 1	0.020 1	0.982 1	0.017 4	0.979 8	0.014 9	0.986 4	0.018 5
WSAE(<i>ABSD</i>)	0.986 6	0.016 6	0.987 0	0.016 5	0.980 4	0.012 2	0.986 3	0.013 9
RAE	0.988 2	0.165 4	0.988 5	0.012 4	0.981 6	0.012 1	0.989 4	0.012 2
WSRAE	0.997 7	0.018 5	0.989 1	0.010 5	0.984 6	0.010 1	0.991 4	0.010 4

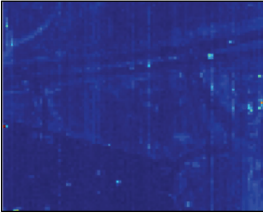
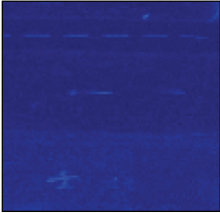
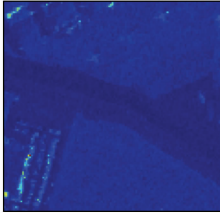
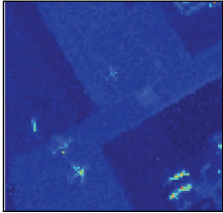
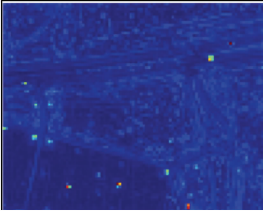
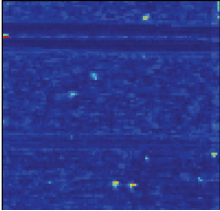
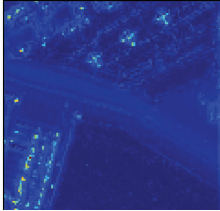
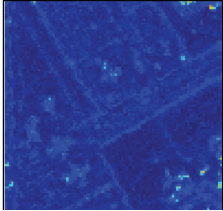
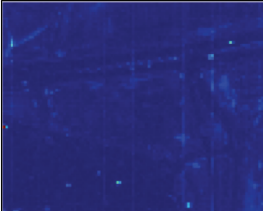
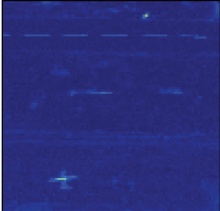
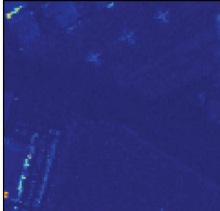
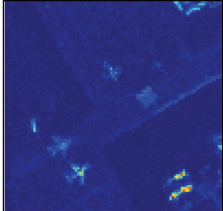
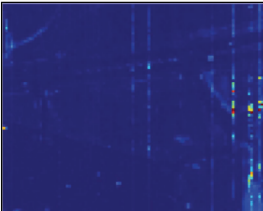
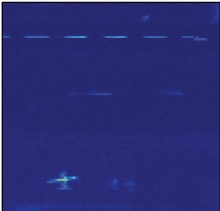
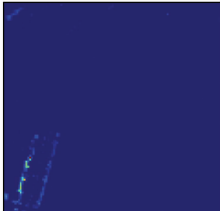
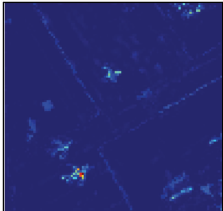
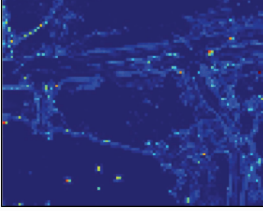
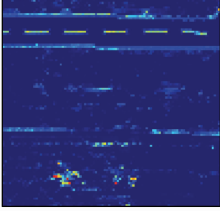
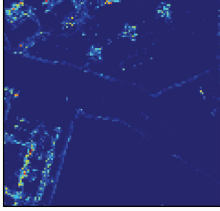
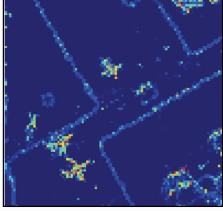
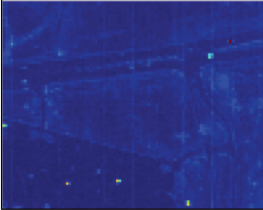
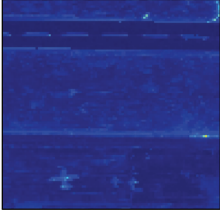
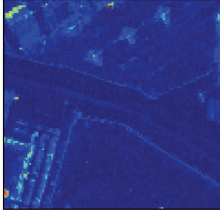
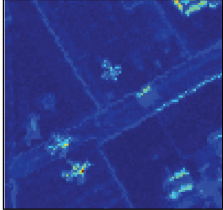
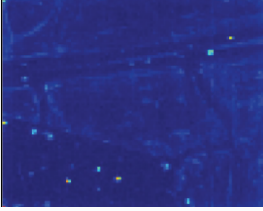
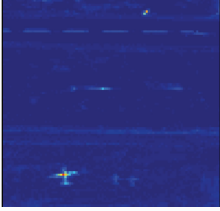
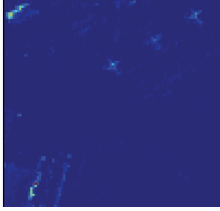
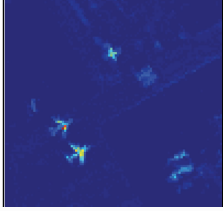
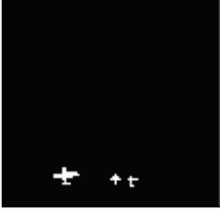
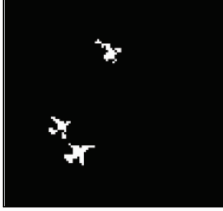
2.3 探测结果与分析

为充分验证本文算法的先进性,本文选取了 6 种先进的异常检测算法进行对比,这 6 种算法分别为 GRX,CRD,LSMAD,FrFE,SC_AAE 和 MemAE^[21] 算法。其中 GRX,CRD,LSMAD 及 FrFE 为经典传统算法,而 SC_AAE 和 MemAE 是先进的深度学习算法。上述深度学习算法在 TensorFlow(v1.10.0,Python 3.6.0 和 CUDA 10.0)中实现,通过使用 Adam 优化网络,初始学习率为 0.000 1,隐藏节点数为 64,批量大小为 1 024,epoch 为 3 000。在本文的 WSRAE 网络中,考虑到网络的像素级处理特性及其对参数的不敏感性,将 MinPts 固定为 1,Eps 设为 0.2,阈值 ε ,设为 0.85。

表 2 展示了 4 组数据集中不同异常检测算法的检测结果对比,本文所提出的 WSRAE 算法获得的检测图最接近异常地物真值图。WSRAE 方法可以确保在抑制丰富类别的背景的同时突出显示具有不同大小和形状的异常,而其他对比算法更有可能忽略异常、被背景分散注意力或失去异常的形状。以

Urban 数据集为例,WSRAE 确保在最小化背景干扰的同时检测到小目标。在 Gulfport 数据集中,WSRAE 可以捕捉不同尺寸的飞机的形状以及它们的位置,但是 GRX,CRD 和 MemAE 并不能有效捕捉。LSMAD 算法几乎失去了 2 个较小的飞机目标。虽然 SC_AAE 也可以清楚地检测到异常,但并不能很好地抑制背景。在 San Diego 数据集中,具有更多结构信息的异常场景中,WSRAE 算法仍然可以很好地检测异常,甚至是目标的边缘。SC_AAE 算法可以发现异常,但它更容易受到背景干扰。如上所述,尽管 GRX 在每个数据集中几乎不受背景的影响,但其异常目标很少被突出显示并且丢失了大部分形状信息。CRD 算法易受窗口大小的影响,无法在所有数据集上获得完整的异常目标。LSMAD 算法并不能完整的检测到异常目标。FrFE 和 SC_AAE 算法在克服噪声和背景方面的鲁棒性较差。MemAE 算法尽管能够探测出异常目标,但其抑制背景性能较差。而本文所提出的 WSRAE 算法在检测异常目标的同时,能够很好地抑制背景。

表 2 4 类数据集不同异常检测算法结果对比

算法	Urban	Gulfport	San Diego - 1	San Diego - 2
GRX				
CRD				
LSMAD				
FrFE				
SC_AAE				
MemAE				
WSRAE				
异常参考图				

本文采用的4组数据集中各异常检测算法的ROC曲线对比如图3所示。从中可以看出,WSRAE算法的ROC曲线总体上位于左上角,显著优于其他对比算法,并且在同等虚警率条件下WSRAE算法能够达到更高探测率。尽管WSRAE算法与其余算

法存在交叉,但在总体上仍位于左上角,仍然可以证明本文方法的先进性。而其余算法在不同数据集中展现的ROC曲线均不稳定,难以适应于所有高光谱数据集。

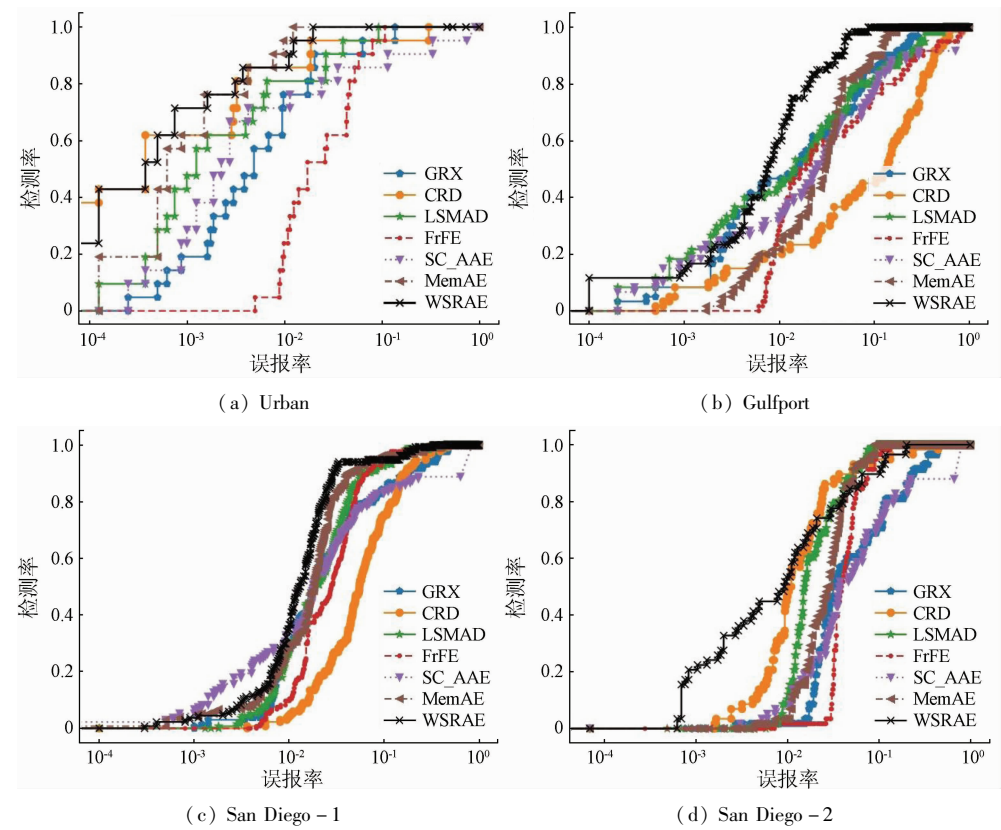


图3 4组数据集中各异常检测算法的ROC曲线对比

Fig.3 Comparison of ROC curves for different anomaly algorithms in 4 data sets

表3对比了各异常检测算法的AUC的(P_d , P_f)和(P_f , τ)值。从表3中AUC(P_d , P_f)值可以看出,WSRAE算法在4组数据集中均获得了最优的检测性能。GRX,CRD和LSMAD算法的探测性能具有不稳定性。尽管SC_AAE与MemAE算法也取得了较高的探测精度,但它们的AUC(P_f , τ)普遍高于WSRAE算法。从AUC(P_f , τ)值可以看出,

WSRAE算法在4组数据集中的误检的概率也是较低的。尽管FrFE算法的AUC(P_f , τ)值在Urban与San Diego - 1中表现为最优,但WSRAE仍然是次优的,仍然可以证明本文方法的优越性。综上所述,表3中的AUC值和图3的ROC曲线与Urban数据集中的异常检测图结果均一致,因此WSRAE方法比其他先进的异常检测算法更具有竞争力。

表3 4组数据集中各异常检测算法的AUC值

检测算法	Urban		Gulfport		San Diego - 1		San Diego - 2	
	(P_d, P_f)		(P_d, P_f)		(P_d, P_f)		(P_d, P_f)	
	(P_d, P_f)	(P_f, τ)	(P_d, P_f)	(P_f, τ)	(P_d, P_f)	(P_f, τ)	(P_d, P_f)	(P_f, τ)
GRX	0.984 8	0.0344	0.952 6	0.024 8	0.911 1	0.040 6	0.940 3	0.058 9
CRD	0.982 9	0.0454	0.821 9	0.054 2	0.971 5	0.051 5	0.919 0	0.063 7
LSMAD	0.989 7	0.0190	0.939 5	0.024 6	0.975 6	0.020 1	0.965 3	0.029 1
FrFE	0.968 7	0.0162	0.894 1	0.049 2	0.953 3	0.002 4	0.962 2	0.021 4
SC_AAE	0.934 8	0.0307	0.893 0	0.024 5	0.861 7	0.019 6	0.890 5	0.028 3
MemAE	0.997 4	0.0337	0.961 6	0.037 9	0.969 7	0.047 0	0.971 3	0.048 7
WSRAE	0.997 7	0.0185	0.986 2	0.015 2	0.9846	0.005 4	0.976 3	0.018 5

表4记录了4组高光谱数据集采用个异常检测算法的计算时间。GRX算法的平均计算时间更快,但其在检测精度中较差。由于局部窗口参数的影响,CRD方法的运行时间比其他方法要长得多。基

于深度学习的方法 SC_AAE, MemAE, WSRAE 比传统方法 CRD, LSMAD 与 FrFE 算法的运算效率更高。WSRAE 算法的时间效率在 3 种深度学习算法中表现最优,这是因为网络上经过训练后直接利用重构误差进行计算异常,而 MemAE 和 WSRAE 包含了相关的后处理步骤,导致运算时间增加。以上分析再次印证了 WSRAE 强大的竞争力。

表 4 异常检测算法的平均计算时间

Tab.4 Average calculation time of each anomaly detection algorithm (s)

算法	Urban	Gulfport	San Diego - 1	San Diego - 2
GRX	0.468	0.445	0.911	0.326
CRD	98.217	45.826	74.654	94.225
LSMAD	26.784	21.941	19.261	25.635
FrFE	20.751	17.462	22.026	17.844
SC_AAE	3.123	5.794	4.145	7.912
MemAE	5.490	5.457	5.404	5.473
WSRAE	2.364	2.455	2.248	2.461

3 结论

本文提出基于弱监督鲁棒性自编码的异常探测算法(WSRAE)。该算法首先提出了一种显著类别搜索方法来搜索样本的伪标签,即弱监督学习所需的不准确标签。为了减轻传统 AE 对噪声和异常值的敏感性,本文引入 $l_{2,1}$ 范数与异常 - 背景光谱距离构建 RAE 框架,以减轻异常部分对参数微调的影响。最后网络生成的重构误差以探测异常目标。采用 4 个高光谱数据集验证了 WSRAE 算法的合理性和优越性,通过将本文算法与其他先进的异常检测算法对比,表明 WSRAE 算法对异常具有更强的鲁棒性,而且能够很好地抑制背景。如何进一步结合空间特征,具体研究异常像元的混合光谱、稀疏性地物问题,将是我们未来工作的重点。

参考文献(References):

[1] 童庆禧,张兵,张立福.中国高光谱遥感的前沿进展[J].遥感学报,2016,20(5):689-707.
Tong Q X,Zhang B,Zhang L F. Current progress of hyperspectral remote sensing in China [J]. Journal of Remote Sensing,2016,20(5):689-707.

[2] 张兵.高光谱图像处理与信息提取前沿[J].遥感学报,2016,20(5):1062-1090.
Zhang B. Advancement of hyperspectral image processing and information extraction[J]. Journal of Remote Sensing,2016,20(5):1062-1090.

[3] 成宝芝.高光谱图像异常目标检测算法研究与进展[J].国土资源遥感,2014,26(3):1-7. doi:10.6046/gtzyyg.2014.03.01.

Cheng B Z. Study and progress of anomaly target detection in hyperspectral imagery[J]. Remote Sensing for Land and Resources, 2014,26(3):1-7. doi:10.6046/gtzyyg.2014.03.01.

[4] 侯增福,刘裕源,闫柏琨,等.基于波段选择与学习字典的高光谱图像异常探测[J].国土资源遥感,2019,31(1):33-41. doi:10.6046/gtzyyg.2019.01.05.

Hou Z F,Liu R Y,Yan B K,et al. Hyperspectral imagery anomaly detection based on band selection and learning dictionary[J]. Remote Sensing for Land and Resources,2019,31(1):33-41. doi:10.6046/gtzyyg.2019.01.05.

[5] Reed I S,Yu X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1990,38(10):1760-1770.

[6] Kwon H,Nasrabad N M. Kernel RX - algorithm: A nonlinear anomaly detector for hyperspectral imagery[J]. IEEE Transactions on Geoscience and Remote Sensing,2005,43(2):388-397.

[7] Guo Q,Zhang B,Ran Q,et al. Weighted-RXD and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2014,7(6):2351-2366.

[8] Du B,Zhang L. Random-selection-based anomaly detector for hyperspectral imagery[J]. IEEE Transactions on Geoscience and Remote sensing,2010,49(5):1578-1589.

[9] Li W,Du Q. Collaborative representation for hyperspectral anomaly detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2014,53(3):1463-1474.

[10] 刘万军,武小杰,曲海成,等.改进协同表示的高光谱图像异常检测算法[J].计算机应用研究,2018(12):68.
Liu W J,Wu X J,Qu H C,et al. Improved collaborative representation for hyperspectral imagery anomaly detection algorithm[J]. Application Research of Computers,2018(12):68.

[11] Zhang Y,Du B,Zhang L,et al. A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2015,54(3):1376-1389.

[12] Tao R,Zhao X,Li W,et al. Hyperspectral anomaly detection by fractional Fourier entropy[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2019,12(12):4920-4929.

[13] Xie W,Liu B,Li Y,et al. Autoencoder and adversarial-learning-based semisupervised background estimation for hyperspectral anomaly detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2020,58(8):5416-5427.

[14] 薛白,王懿哲,刘书含,等.基于孪生注意力网络的高分辨率遥感影像变化检测[J].自然资源遥感,2022,34(1):61-66. doi:10.6046/zrzyyg.2021122.
Xue B,Wang Y Z,Liu S H,et al. Change detection of high-resolution remote sensing images based on Siamese network[J]. Remote Sensing for Land and Resources,2022,34(1):61-66. doi:10.6046/zrzyyg.2021122.

[15] 韩彦岭,崔鹏霞,杨树瑚,等.基于残差网络特征融合的高光谱图像分类[J].国土资源遥感,2021,33(2):11-19. doi:10.6046/gtzyyg.2020209.

Han Y L, Cui P X, Yang S H, et al. Classification of hyperspectral images based on feature fusion of residual network [J]. Remote Sensing for Land and Resources, 2021, 33 (2) : 11 – 19. doi: 10.6046/gtzyyg. 2020209.

[16] Xie W, Lei J, Liu B, et al. Spectral constraint adversarial autoencoders approach to feature representation in hyperspectral anomaly detection [J]. Neural Networks, 2019, 119 : 222 – 234.

[17] Lu X, Zhang W, Huang J. Exploiting embedding manifold of autoencoders for hyperspectral anomaly detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 58 (3) : 1527 – 1537.

[18] Zhao C, Zhang L. Spectral – spatial stacked autoencoders based on low – rank and sparse matrix decomposition for hyperspectral anomaly detection [J]. Infrared Physics & Technology, 2018, 92 : 166 – 176.

[19] Parimala M, Lopez D, Senthilkumar N C. A survey on density based clustering algorithms for mining large spatial databases [J]. International Journal of Advanced Science and Technology, 2011, 31 (1) : 59 – 66.

[20] Ma Y, Li C, Mei X, et al. Robust sparse hyperspectral unmixing with $l_{2,1}$ norm [J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 55 (3) : 1227 – 1239.

[21] Gong D, Liu L, Le V, et al. Memorizing normality to detect anomaly: Memory – augmented deep autoencoder for unsupervised anomaly detection [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019 : 1705 – 1714.

Hyperspectral anomaly detection based on the weakly supervised robust autoencoder

ZHANG Guojian¹, LIU Shengzhen², SUN Yingjun¹, YU Kaijie³, LIU Lina⁴

(1. College of Surveying and Geo – Informatics, Shandong Jianzhu University, Jinan 250101, China; 2. First Geodetic Survey Team, Ministry of Natural Resources, Xi’an 710054, China; 3. Zhejiang Zhongcexintu Geographic Information Technology Co., Huzhou 313200, China; 4. Ningbo Yinzhou District Surveying and Mapping Institute, Ningbo 315041, China)

Abstract: Hyperspectral anomaly detection has received particular attention due to its unsupervised detection of targets. Moreover, autoencoder (AE), together with its variants, can automatically extract deep features and detect anomalous targets. However, AE is highly generalizable due to the existence of anomalies in the training set, thus suffering a reduced ability to distinguish anomalies from the background. This study proposed an anomaly detection algorithm based on the weakly supervised robust AE (WSRAE). First, this study developed a salient category search strategy and used probability – based category thresholds to label coarse samples in order to make preparation for network – based weakly supervised learning. Moreover, this study constructed a robust AE framework constrained jointly by $l_{2,1}$ norm and anomaly – background spectral distances. This framework was robust with regard to noise and anomalies during training. Finally, this study detected anomalous targets based on the reconstruction errors obtained from all samples. Experiments on four hyperspectral datasets show that the WSRAE algorithm has greater detection performance than other state – of – the – art anomaly detection algorithms.

Keywords: hyperspectral image; anomaly detection; salient category search; robust AE

(责任编辑: 李 瑜)