

DOI: 10.20189/j.cnki.CN/61-1527/E.202603001

高光谱目标检测：从传统方法到人工智能的演化

张志利, 孙海翔*, 丁 遥, 鲁振航, 李怀淳, 李 健, 吴俊豪
(火箭军工程大学, 西安 710025)

摘 要: 为系统剖析高光谱目标检测 (hyperspectral target detection, HTD) 技术由传统数学建模向深度学习框架演进的发展脉络, 定量评估了各类代表性算法的性能表现。在研究方法上, 横向对比了约束能量最小化与自适应匹配滤波等经典模型, 深入解析了三大前沿智能机制的核心创新: 一是将卷积神经网络 (convolutional neural network, CNN) 作为基于空间-光谱联合表征的非线性关系度量器; 二是利用 Transformer 的多头自注意力机制结合生成式自监督范式, 准确捕获全局长程光谱依赖; 三是依托自编码器潜在空间中的多维物理约束, 实现微弱目标与复杂背景的深度解耦。基于 Sandiego1、Sandiego2 与 Cuprite 数据集, 以特征曲线下面积 (area under curve, AUC) 为核心评价指标, 对 10 种不同算法开展了实验验证。结果表明: 受制于单像素特征拟合与高维冗余, 传统方法的 AUC 值均普遍难以突破 0.81; 相反, 深度学习模型的 AUC 则全面跃升至 0.86 以上。具体而言, CNN 架构的 SCL-HTD 模型凭借卓越的长距离相关性捕获能力, 在背景高度混杂的 Sandiego2 数据集上的 AUC 取到最大值 0.897 4; 基于 CNN 的 S2ADet 算法在空间目标分布规则的 Cuprite 数据集中 AUC 值达到 0.894 5; 而融入无监督重建机制的自编码类模型 NUN-UTD 同样在 Sandiego1 中取得了 AUC 值高达 0.895 2 的最佳表现。由于单一网络模型已触及特征表征的瓶颈, 未来高光谱探测技术的破局点在于构建自适应混合架构及多粒度特征融合机制, 并向无监督与弱监督范式全面倾斜, 以此来攻克复杂环境泛化与小样本特征迁移的问题。

关键词: 高光谱图像; 目标检测; 深度学习; 图像处理

中图分类号: TP751.1

文献标志码: A

开放科学标识码(OSID):

文章编号: 2097-2741(2026)03-0001-23



听语音
与作者
聊科研

近几十年来, 科学技术的进步使得高光谱成像在光学遥感领域内变得愈发重要。从光谱角度来看, 高光谱图像通常包含数百个窄带光谱通道, 可覆盖从可见光到近红外, 甚至中红外波段的完整电磁波谱^[1]。相比于多光谱图像来说, 高光谱图像的光谱维度更高, 可以很好地体现出不同目标之间的微小波段差别, 在城市规划^[2]、土壤制图^[3]和灾害管理^[4]等领域应用广泛。高光谱目标检测 (hyperspectral target detection, HTD) 本质上是一个二分类问题, 即用已知目标的先验光谱特征去

识别高光谱图像中特定的目标, 主要的难点在于辨别出目标像素和背景像素。

需要指出的是, 本文描述的“人工智能”, 主要是指深度学习在 HTD 中的技术范式。尽管人工智能的范围更为广泛, 但近十年来推动 HTD 性能取得突破性进展的核心驱动是深度神经网络。因此, 本文内容聚焦于深度学习框架下的代表性模型, 系统分析其对 HTD 的技术影响, 主要工作如下。一是系统性梳理与对比: 首次在同一实验框架下, 横向对比了经典统计模型和前沿深度学习

收稿日期: 2026-04-07

修回日期: 2026-05-07

录用日期: 2026-05-29

第一作者: 张志利 (1966—), 男, 教授, 主要从事发射平台定位定向与导航技术研究。

***通信作者:** 孙海翔 (2003—), 男, 硕士研究生, 主要从事高光谱图像目标检测研究。

引用格式: 张志利, 孙海翔, 丁遥, 等. 高光谱目标检测: 从传统方法到人工智能的演化[J]. 火箭军工程大学学报, 2026, 40 (3): 1-23. ZHANG Zhili, SUN Haixiang, DING Yao, et al. Hyperspectral target detection: the evolution from conventional to artificial intelligence[J]. Journal of Rocket Force University of Engineering, 2026, 40 (3): 1-23.

(三大类共 10 种代表性算法) 的性能表现; 二是定量验证技术演进: 通过计算特征曲线下面积 (area under curve, AUC) 指标, 在 3 个标准数据集上定量揭示了从人工特征到数据驱动的技术迭代所带来的性能跃升; 三是深入剖析架构-数据适配性: 研究了不同深度学习架构与特定数据集内在特性间存在的强适配关系, 为后续算法选型与设计提供了实证依据。

1 高光谱图像目标检测的研究现状

早在 20 世纪 80 年代初, 研究人员就凭借信号处理和统计学基础理论, 通过建立数学模型来捕获目标和背景的频谱差异, 进而将目标从背景中分离出来。这一时期形成了子空间投影、约束能量最小化和假设检验等重要的理论基础。1994 年, Harsanyi 等^[5]提出了正交子空间投影 (orthogonal subspace projection, OSP) 法, 将高光谱数据映射到背景子空间的正交补集中, 开创了子空间类方法的先河。随后, Harsanyi 又提出了一种约束能量最小化 (constrained energy minimization, CEM) 方法^[6], 以最大化目标能量和最小化背景能量为约束条件来优化模型, 可有效获取目标响应值。在此基础上, 学术界又发展出多目标 CEM^[7]和集成 CEM 技术^[8]。同时, 基于二元假设检验框架以及广义似然比检验理论, 又相继提出自适应匹配滤波器 (adaptive matched filter, AMF)^[9]和自适应余弦估计器 (adaptive cosine estimator, ACE)^[10]等。从 20 世纪 90 年代到 21 世纪初, “如何在强背景干扰下稳定提取目标响应” 是高光谱目标检测的主要任务。以 CEM、AMF、ACE 为代表的方法, 虽然形式不同, 但其主要目的都是 “背景抑制-目标增强”, 为后续高光谱目标检测研究奠定了理论基础。特别是在亚像素目标、先验目标已知而标注样本稀缺的应用情境下, 这类方法一度成为高光谱目标检测的标准工具。

伴随着机器学习技术的发展, 传统线性假设的局限性也逐渐凸显出来。为得到更多的非线性结构信息以提高检测精度, 核方法、稀疏表示、度量学习等新技术也逐渐被引入 HTD 领域中。核方法通过把输入数据映射到高维特征空间中, 成功将线性子空间模型推广到了非线性领域, 其中典型的方法为核约束能量最小化^[11]。随着核方法与再生核希尔伯特空间理论逐步成熟, 研究人员开始着手打破传统线性探测器的固有限制。2005

年, Kwon 等^[12]提出核正交子空间投影, 次年又提出核自适应子空间检测器^[13], 通过核映射将原始空间中难以进行线性分离的目标-背景关系转化为高维特征空间中的可分问题, 为后续研究提供了参考。除此之外, 稀疏表示理论的出现又引起了 HTD 技术领域的重大突破。2011 年, Chen 等^[14]首次提出了使用稀疏表示进行目标检测的方法; 之后, Li 等^[15]将稀疏表示和协同表示结合起来, 进一步提高了检测性能。2014 年, Zhang 等^[16]将有监督度量学习融入到 HTD 研究中, 创建出新的距离度量准则, 使目标和背景的特征差异得到最大程度的体现, 从而提高了检测准确率。

深度学习在 HTD 中的应用经历了一个循序渐进的发展过程。2017 年, Li 等^[17]首次尝试将迁移学习用于 HTD 领域, 通过构造像素对与差分特征来解决训练样本不足的问题, 从而反映出深度学习在 HTD 中的可行性。2018 年, Du 等^[18]进一步提出基于减法像素对的 CNN 检测框架, 为利用深度学习进行 HTD 研究提供了重要启发。近年来, 以卷积神经网络、自动编码器 (autoencoder, AE)、生成对抗网络 (generative adversarial network, GAN)、图神经网络 (graph neural network, GNN)、Transformer 等为代表的深度学习模型, 凭借其能够从高维光谱数据中自动提取深层次特征的显著优势, 大大增强了检测算法的准确度和鲁棒性。2021 年, Zhu 等^[19]提出一种用于 HTD 的双通道卷积网络; 同年, Xie 等^[20]探索了无监督 GAN 在 HTD 中的应用; 2024 年, Xu 等^[21]进一步提出了结合 GNN 与 CNN 的联合框架, 推动了 HTD 技术的持续演进。与判别式网络不同, AE 更强调通过背景重构、潜变量约束和异常分离来实现目标检测。2020 年, Xie 等^[22]提出光谱正则化无监督网络, 通过增强背景的重构一致性, 显著凸显了目标的异常响应。2022 年, 研究人员开始研究层次去噪、自监督聚类、变分编码及非线性解混问题^[23-25], 使 AE 逐渐演变为兼具背景建模、目标增强与物理先验融合能力的检测框架。与此同时, 对深度学习的研究也逐步呈现出 “结构化” 和 “融合化” 的特点。深度神经网络开始重新尝试使用传统的检测原理, 例如 Yang 等^[26]提出的深度约束能量最小化 (deep CEM, DCEM) 方法, 将约束能量最小化准则直接融入神经网络来进行目标优化。同时, GCN、Transformer 以及一些其他深度神经网络的引入, 使得深度学习不再只是简单地作为特征提取器, 而是逐步成为融合物理

先验、全局依赖建模与可解释优化过程的统一平台^[27]。

综上所述, HTD 的研究现状具有明显的阶段性特征。第一阶段, 以 CEM、AMF、ACE 等为代表, 注重加强对统计推导和可解性的研究; 第二阶段, 以核方法、稀疏表示和度量学习为代表, 致力于突破线性假设并增强特征可分性; 第三阶段, 以 CNN、AE、GAN、Transformer 及其融合模型为代表, 注重深度学习在 HTD 中的使用, 并逐步对其进行完善。传统方法提供了可解释的检测原理, 机器学习方法扩展了非线性表征能力, 而深度学习方法则在大规模特征学习和复杂环境泛化方面表现出色。然而, HTD 面临以下三大核心挑战。一是维度灾难与信息冗余: 数百个连续波段带来高维数据处理难题, 其中包含大量冗余与噪声, 增加了计算复杂度并可能降低检测性能; 二是复杂背景干扰: 真实场景中背景呈现多模态、非高斯分布特性, 且与目标光谱高度相似, 导致

传统的基于统计假设的方法失效; 三是样本稀缺与泛化难题: 高质量像素级标注成本高昂, 导致监督学习方法受限, 模型在跨传感器、跨场景应用时泛化能力不足。

针对上述挑战, 研究人员已形成若干共识。首先, 单纯依赖单像素光谱信息已无法满足需求, 空间-光谱联合建模成为提升性能的关键路径; 其次, 从依赖人工设计特征的传统方法, 向数据驱动的端到端深度学习范式演进是必然趋势; 最后, 未来研究需向弱监督、无监督及自监督学习倾斜, 以解决标签数据匮乏问题, 并突破单一模型的表征瓶颈。遵循上述技术发展脉络, 后续章节将围绕传统统计模型、机器学习方法与深度学习框架下的代表性 HTD 算法展开系统梳理, 并对各类代表性方法展开详尽的实验对比与性能评估, 在全面剖析实验结果的基础上, 指明高光谱目标检测技术未来的潜在研究方向。图 1 所示为高光谱图像目标检测方法发展脉络图。

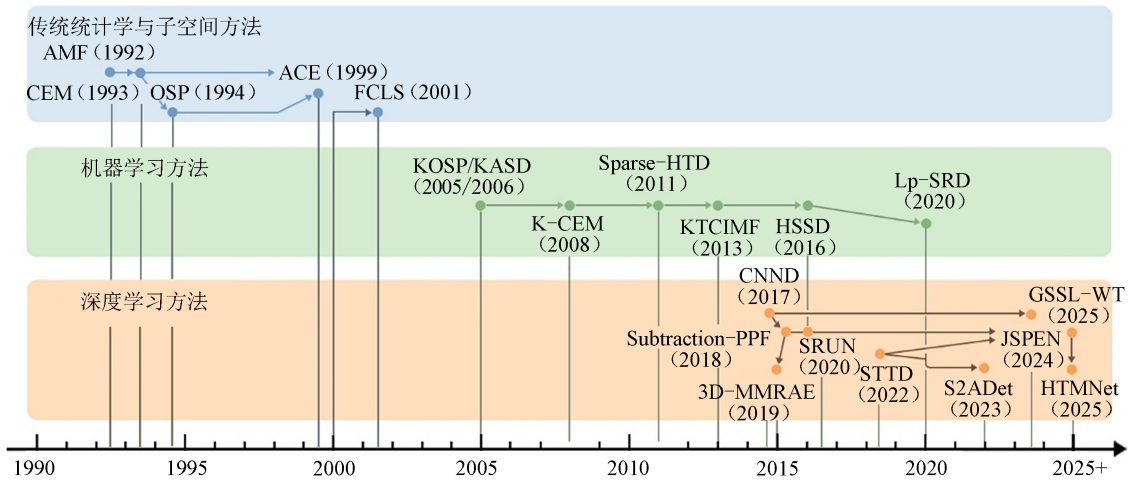


图 1 高光谱图像目标检测方法发展脉络

Fig. 1 Development trajectory of hyperspectral target detection methods

2 传统的高光谱图像目标检测方法

2.1 基于结构化的背景模型

在结构化背景模型的假设下, 图像中任意像素的光谱特征都可通过若干基元光谱的线性叠加来表示。其中, 对应基元的权重系数称为丰度^[28]。

2.1.1 有约束检测

有约束检测的关键在于为基元丰度施加了一个符合实际物理意义的约束条件。针对无约束检测方法在求解时易出现负丰度的问题, Chang 等^[29]于 2000 年提出了非负约束最小二乘法 (nonnegative constrained least square, NCLS), 以最小二乘误差作为优化目标, 通过加入非负性约束条件,

将无约束求解出来的负丰度分量截断修正为 0, 虽然并不能保证所有的丰度分量之和为 1, 但却消除了负值所带来的计算干扰, 提高了目标的可检测性。为进一步契合传统的线性光谱混合模型, 即丰度必须同时满足非负且总和为 1 的真实物理属性, Heinz 等^[30]在 2001 年进一步提出了全约束最小二乘法 (fully constrained least square, FCLS), 在优化过程中严格落实“丰度和为一约束” (abundance sum-to-one constraint, ASC) 与“非负性约束” (abundance non-negativity constraint, ANC)。在具体实现上, FCLS 把最小均方误差作为优化目标, 通过构建扩展端元矩阵与扩展像素向量, 将 ASC 条件无缝融入现有模型中, 最后依靠 NCLS

算法的迭代优化框架,完成了对 ANC 与 ASC 约束条件的联合求解。

2.1.2 无约束检测

无约束检测方法不再强调丰度必须严格满足真实物理条件,而是直接通过滤波、投影或能量约束来实现目标增强与背景抑制。其中,CEM 是最具代表性的方法之一,其核心思想是在保持目标光谱响应的同时,尽可能降低背景输出能量,可以被理解作为一种典型的“目标通过、背景抑制”检测器。与有约束丰度估计方法相比,CEM 结构更加简洁,计算实现更为直接,也更适合作为传统 HTD 的基准方法。

针对传统 HTD 方法存在残留背景干扰的问题, Yang 等^[31]率先在 CEM 中引入 L1 范数正则化项,通过二次锥规划和内点法来对模型进行高效求解,显著抑制了背景干扰并降低了误报率。在结构优化方面, Zou 等^[32]设计了一种多级级联的 CEM 检测架构,在每一级检测后利用非线性抑制函数对光谱向量进行加权更新,并将其输入到下一级中继续迭代直至收敛,这种逐级削弱背景振幅的机制使得深层探测器能够更精准地聚焦于隐蔽目标。此外, Zhao 等^[8]提出多级级联的集成极限学习机模型,通过级联非线性变换来增强特征判别力,同时使用随机平均策略来大幅提升模型泛化能力。针对高光谱数据存在维度灾难与信息冗余的问题, Shang 等^[33]把目标约束干扰最小化滤波器的思想应用到频段选择上,提出一种目标约束抗干扰波段选择方法 TCIMF,能够面向特定目标检测优选波段子集,从而抑制无关目标和背景干扰。为克服先验目标光谱误差所造成的模型检测性能衰退问题, Zhang 等^[34]提出了一种鲁棒的贝叶斯约束能量最小化算法,摒弃了传统的目标光谱单点估计方式,依托先验目标光谱推导真实目标光谱的后验分布。在变换域增强方面, Zhao 等^[35]提出了基于分数阶傅里叶变换和多方向双窗口策略的自适应方法,通过分数阶变换有效地放大了目标和背景的分离度,进一步改善了经典的能量最小化框架。在底层矩阵运算和背景抑制机制的改进上, Fu 等^[36]认为传统的 CEM 在计算自相关矩阵时若混入目标像素,则会导致目标自身受到抑制,因此提出了相关改进算法来避免此类问题的发生,从而提高了检测精度。Zhang 等^[37]从算法稳定性角度出发,提出了正则化 CEM 算法。2020 年, Shang 等^[38]对目标约束干扰最小化波段选择(target-constrained interference-minimized band

selection, TCIMBS) 框架进行了进一步完善,筛选出的波段组合既能实现对目标波段特征的高效表征,又可抑制无关干扰目标和地物背景;2021 年, Sun 等^[39]将子空间划分和 CEM 方法相结合,先将频段划分成若干个弱相关子集,然后根据 CEM 方差准则选出最具区分度的频段。为应对极为复杂的背景干扰, Chen 等^[40]在 2022 年提出了一种基于先清理后检测的思想来对 TCIMF 方法进行改进,该方法首先去除全局背景,再利用 CEM 处理筛选后的数据,从而提高复杂环境下的目标检测效果。2022 年, Tian 等^[41]针对 CEM 目标检测提出了一种新的波段选择方法,以最小化输出能量为准则,结合稀疏约束找到最佳波段组合,显著提升了波段和探测器之间的匹配效率。

2.2 基于非结构化的背景模型

基于非结构化背景模型的方法通常不再显式构造端元或背景子空间,而是将目标检测转化为统计假设检验问题。该类方法假设待测像元可能属于背景假设或目标假设,并通过背景统计量、目标先验光谱和相似性度量构造检测统计量。其中, AMF 和 ACE 是最具代表性的 2 类方法。

AMF 通过估计背景协方差实现自适应匹配滤波,适用于背景统计特征相对稳定且目标先验光谱已知的场景,其优势在于具有明确的统计检测意义。ACE 则更强调待测像元与目标光谱之间的角度一致性,相较于单纯幅值匹配,对光照强度变化和部分幅值扰动具有更强鲁棒性。然而,上述 2 种方法的有效性高度依赖背景统计量的估计质量。当背景呈现多模态、非高斯或强非平稳特征时,协方差估计可能不准确,进而导致误报率上升。同时,基于非结构化背景模型的方法通常以单像素光谱为主要判别依据,难以利用空间结构和邻域上下文信息,因此在复杂真实场景中往往需要与鲁棒估计、局部背景建模或深度特征学习方法结合。Wang 等^[42]提出一种迭代加权的 ACE 算法来处理由传感器误差、大气散射和材料物理性质不同而引起的光谱变化,可在训练数据有限的情况下得到质量较好的目标参考光谱,从而提高了检测的鲁棒性。针对训练数据标签不准、目标像素占比较少的实际问题, Zare 等^[43]提出了 2 种新的多实例目标特征化手段,大幅扩展了基于非结构化背景模型的适用范围。

2.3 基于机器学习的方法

2.3.1 核方法

核方法的引入主要是为了解决传统线性检测

器难以处理非线性光谱混合的问题, 其基本思想是通过核映射将原始光谱数据隐式投影到高维特征空间, 从而使原本难以线性分离的目标与背景获得更好的可分性。但核方法的性能高度依赖核函数类型和核参数选择, 且核矩阵会带来较高的计算复杂度, 不利于对大规模高光谱图像的快速处理。

为了解决传统线性模型因波段少、判别信息不足而导致模型检测精度较低的问题, Jiao 等^[11]于 2008 年首次将高斯径向基函数 (radial basis function, RBF) 应用到约束能量最小化框架中, 提出核约束能量最小化方法 K-CEM, 利用核映射把数据投影到非线性特征空间中, 从而提高了模型对混合像素以及亚像素目标的检测效果。Wang 等^[44]针对高光谱亚像素检测中非线性端元混合和目标背景高度相似的干扰问题, 提出核目标约束干扰最小化滤波器 KTCIMF, 对目标和干扰子空间进行了双重约束, 不仅可以找到目标, 还可以有效抑制背景杂波和干扰信号。为解决高光谱图像因光谱特征重叠而导致类别难以区分的问题, Liu 等^[45]提出核谱角匹配算法, 将传统的谱角计算转换至再生核希尔伯特空间中, 并对多项式、Sigmoid、高斯 RBF 等不同的核函数进行了系统评价。此外, 在自适应子空间检测领域, 核方法同样展现出解决非线性难题的巨大潜力。Kwon 等^[13]利用高斯 RBF 核, 提出了核自适应子空间检测器 KASD, 通过核机制挖掘谱带间的高阶统计相关性, 成功克服了传统自适应子空间检测器通常受限于二阶统计信息且难以适应非线性光谱变化的固有缺陷, 在背景噪声协方差未知或估计不准的环境中表现出卓越的鲁棒性。

2.3.2 稀疏协同表示

稀疏表示与协同表示的基本思想是: 若待测像元能够在背景字典中较好重构, 则更可能属于背景; 若其与目标字典更匹配, 或是在背景字典下产生较大重构残差, 则可能为目标。与直接光谱匹配相比, 稀疏表示方法能够利用字典结构表达目标与背景之间的组合关系, 对混合像元和含噪像元具有一定的鲁棒性。协同表示方法则在一定程度上降低了严格稀疏约束所带来的计算复杂度, 提高了检测效率。但此类方法的性能受样本代表性、字典纯度和计算复杂度的影响显著。

针对高光谱数据中存在的目标先验概率低、样本严重不平衡等问题, Du 等^[46]提出了混合稀疏和统计检测器, 将稀疏表示在高维特征提取上的

优势同统计模型在决策制定上的鲁棒性相融合, 通过特定字典设计、背景字典净化、广义似然比检验等技术, 明显提升了模型对小目标、亚像素目标的检测性能。在稀疏表示理论方面, Chen 等^[14]提出一种基于稀疏表示的高光谱自动目标检测算法, 认为高光谱像素分布在一个低维子空间中, 可通过训练样本的稀疏线性组合来表示, 进而利用压缩感知理论求解得到测试样本的稀疏特征。随后, Chen 等^[47]提出基于非线性核的目标检测算法, 通过在特征空间中创建核化联合稀疏模型, 加强了相邻像素之间的空间相关性, 提升了模型在复杂环境下的检测稳定性。随着稀疏理论的发展, Shi 等^[48]将复杂的数学结构加入到背景抑制和目标分离中, 提出了一种双线性稀疏目标检测器, 将双线性稀疏混合模型无缝融合至自适应匹配子空间检测器, 有效剔除了高纯度、复杂的背景样本, 极大地提高了背景子空间的可信度。针对传统的稀疏检测器对信号表示差和检测残差难以区分比较且对亚像素目标检测能力不足的问题, Zhang 等^[49]提出了基于稀疏表示的二元假设检测模型, 在此基础上又提出了核化二元假设检测器^[50], 并将其推广到非线性空间, 使检测器的性能实现了跨越式发展。2019 年, Ling 等^[51]提出一种约束稀疏表示的双假设目标检测模型, 通过非负约束和自适应上限约束来减小背景字典中目标干扰的影响, 并结合残差差异和稀疏权重来提高目标与背景的可分性。2020 年, Zhao 等^[52]从范数优化的角度出发, 提出了基于自适应迭代收缩阈值法的 LP 范数稀疏表示模型, 通过在统一框架下引入非凸 LP 范数约束, 并结合字典自适应来优化特征的稀疏度和检测精度。2022 年, Zeng 等^[53]提出基于稀疏张量模型的光谱角检测器, 利用三维张量分解技术剥离稀疏目标成分, 并结合光谱角的子空间映射和统计来优化策略。进入深度学习时代, Dong 等^[27]于 2024 年提出了一种可解释的高光谱目标检测网络 JSPEN, 将自适应空间-光谱联合稀疏模型的迭代优化过程嵌入进端到端可训练的深度框架中, 通过稀疏表示理论给网络赋予了物理可解释性, 有效解决了传统深度学习的“黑箱”问题, 使网络可自适应地提取高光谱图像中的背景和目标稀疏特征。

2.4 传统方法的问题与局限性

上述传统方法为高光谱图像目标探测奠定了理论和技术基础, 但在实际应用过程中仍然存在以下诸多短板和不足。

1) 高光谱数据中固有的数百个连续波段易导致“维度灾难”，不但大幅提升了计算复杂度和存储成本，而且混杂其中的低信噪比波段和冗余信息还会引入额外噪声，从而影响模型整体的检测精度。

2) 传统方法在底层逻辑上一般均假设背景像素服从多元高斯分布，但在真实的复杂场景下，异质背景的多模态、非对称统计特征却极为复杂，使得传统方法的最终检测效果无法满足实际需求。

3) 传统方法对复杂环境的适应能力明显不足。受到目标自身材质、光照条件突变、大气物理效应等因素影响，目标光谱极易发生变化，且由于设备空间分辨率不够，易出现“混合像素”，大大增加了高光谱图像目标检测的误报率和漏检率。更严重的是，传统模型参数大多仅针对某个特定场景进行设定，导致其对于不同成像条件缺少必要的泛化能力。

4) 多数传统算法仅局限于挖掘单一像素的光谱特征，几乎完全忽略了目标在空间上的连续性

与结构一致性。不仅如此，现有的处理流程通常被割裂为预处理、特征降维和目标检测等多个独立环节，缺乏端到端的统一优化机制，从而限制了其整体性能的进一步突破。

3 基于深度学习教育的HTD方法

3.1 卷积神经网络

3.1.1 基于样本对匹配的深度学习度量与判别机制

基于CNN的高光谱图像目标检测方法并不直接判断单个像素是否为目标，而是通过构造“目标-待测像元”或“背景-待测像元”的样本对，让网络学习二者之间的相似性关系。根据输入变量的不同，该方法可分为2种模式：一是显式差分模式，即先对输入样本对的光谱特征做差，再将其输入到CNN模型中进行判别；二是双分支匹配模式，即分别提取目标或背景的深层特征，再在嵌入空间中计算其与待测像元的相似性。基于CNN的高光谱图像目标检测2种模式原理如图2所示。

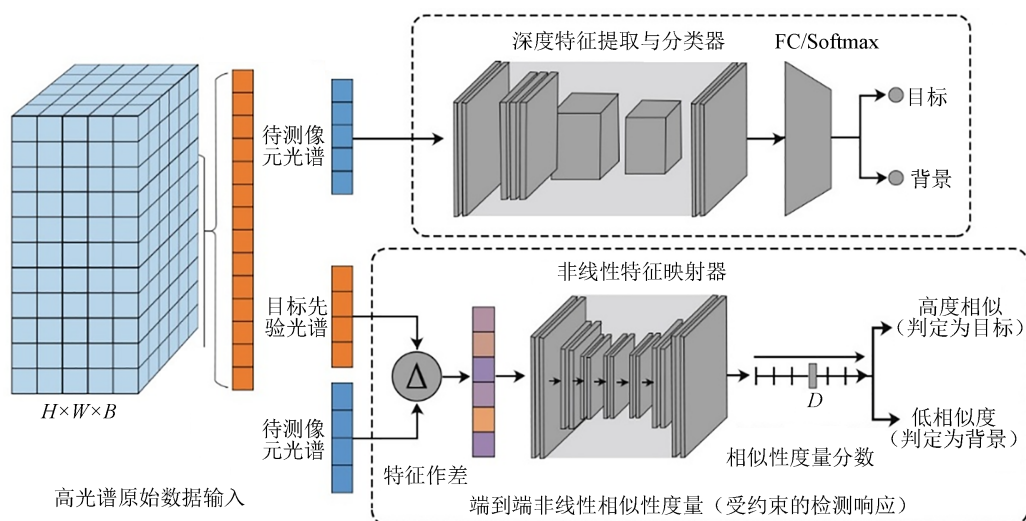


图2 基于CNN的高光谱目标检测双模式原理

Fig. 2 Dual-mode principle for CNN-based hyperspectral target detection

从图2可以看出，基于CNN的高光谱图像目标检测方法的核心贡献不是引入复杂卷积结构，而是从根本上改变了高光谱图像目标检测的判别单位，即检测对象由“单像元类别”转变为“样本对关系”。这一转变缓解了目标样本稀缺的问题，研究人员可以通过有限的先验光谱配对产生大量训练样本，但显式差分模式可能损失原始光谱细节，双分支结构对训练复杂度和样本构造敏感性更高。

通过不断创新网络输入与度量机制，研究者

逐步解决了样本稀缺与特征表征难题，形成了一条包含六项关键进展的清晰发展脉络。2017年，Li等^[17]率先提出基于迁移学习的深度CNN框架，通过构造像素对差分来训练多层网络，打破了CNN在高光谱图像中不能直接进行类别推断的固有局限，首次证明了CNN网络可以依靠学习像素间的相似关系来实现目标检测。2018年，Du等^[18]提出减法思想，通过将目标和背景的差分作为正样本、背景与背景间的差分作为负样本解决了因目标样本不足所导致的训练规模受限问题。同年，

Du 等^[54]对上述思想进行了系统扩展，正式提出 subtraction-PPF 框架，通过目标类来完善样本组合机制，标志着像素对扩增、相似性判别、邻域聚合路线的成熟。然而，单纯的样本扩增和 0/1 硬标签不能应对复杂的高光谱目标检测。因此，Zhang 等^[55]于 2020 年提出 HTD-Net，通过改进的自编码器生成潜在目标，并结合线性预测筛选背景，首次提出了连续相似度回归机制，使网络从二分类走向了更精细的“相似程度”区分。同年，为了克服输入层显式差分表征不足的问题，Feng 等^[56]提出一种基于孪生网络的谱空联合检测模型，

利用双分支网络提取特征，通过欧氏距离和对比损失把匹配关系从原始数据空间提升到深层次的非线性特征空间。针对前期差分操作所导致的判别性光谱细节丢失的问题，Zhu 等^[19]于 2021 年提出了双流卷积网络，直接以目标先验和待匹配像素的原始光谱对作为输入，采用比例混合的像元合成法模拟真实背景，并通过 ICS 损失函数显著拉大了类间间隔。至此，该研究方向完成了从输入端显式差分驱动到特征域深度度量学习，再到原始光谱双流匹配的跃迁。各种方法的优缺点如表 1 所示。

表 1 基于样本对匹配的深维度量与判别机制 CNN 方法对比

Table 1 Comparison of CNN methods with deep metric and discriminative mechanism based on sample-pair matching

方法	年份	优势	局限性
CNND ^[17]	2017	像素对可扩充训练样本	减法会损失光谱细节，高度依赖标注数据
Subtraction Pixel Pair-Based CNN ^[18]	2018	通过减法构造二值(0/1)样本，并采用 CNN 提取判别特征	硬标签不够精细，减法会丢失细粒度信息
Subtraction-PPF ^[54]	2018	像素对构造框架稳健，检测精度高	计算时间长，减法导致信息损失
HTD-Net ^[55]	2020	自编码器生成逼真目标样本，并采用细粒度回归度量	输入差分会降低判别能力，背景筛选不完全
S-CNN ^[56]	2020	孪生网络避免减法，并通过后处理利用空间上下文	依赖初始检测结果，需要源域数据
TSCNTD ^[19]	2021	原始光谱对保留细节，合成样本扩大类间距离	结构复杂，训练极其耗时

3.1.2 空间-光谱联合深度表征网络

空间-光谱联合深度表征网络通常是将中心像素及其邻域构造成局部三维数据块，通过 3D 卷积、双分支网络或空间-光谱特征融合模块同步获取目标和背景的光谱差异和空间上下文信息。典型的融合跨域迁移与空谱特征联合提取的 CNN 深度高光谱目标检测网络如图 3 所示。

结合图 3 可以看出，CNN 在该类方法中主要充当“联合表征器”的角色。光谱分支用于保持材料识别所需的波段信息，空间分支用于补充邻域纹理、目标形态和局部上下文信息，融合模块则负责将二者转化为可判别的目标响应。与样本对匹配方法相比，空间-光谱联合表征网络更适合目标具有一定空间连续性或局部结构稳定的场景，但该方法对训练样本、窗口尺度和空间分辨率更为敏感，在对小目标或亚像素目标的检测中可能受到邻域背景影响。

从发展脉络看，此类 CNN 方法经历了由局部空谱联合建模到特征解耦，再到场景级定位和迁

移泛化的过程。早期的研究主要验证网络如何重塑目标和复杂背景之间的关系，以及边界如何识别。Freitas 等^[57]针对海上监视任务率先提出 3D CNN，通过对局部空间-光谱数据块进行同步卷积运算，提取得到面向目标识别的耦合特征，为联合表征研究奠定了理论基础。随着研究的深入，Li 等^[58]打破传统卷积模式，采用双分支残差网络，将空间邻域和光谱线索分开表征，提出中心损失来压缩类内离散度，使 CNN 摆脱了端到端分类器的束缚，成为专门优化表征空间紧致度的特征嵌入器。He 等^[59]在双流结构和特征聚合模块的基础上创建起一个完整的检测网络，并在自建的 HOD3K 数据集上验证了该方法的有效性，从而将 CNN 模型从“小样本局部特征识别”发展为“完整目标框定位系统”。针对复杂的应用环境，Zhao 等^[60]对地基观测中经常出现的目标谱漂移、背景扰动等工程问题进行了研究，通过在深度网络前引入多尺度谱相似度度量的方式，结合 3D-2D CNN 级联以及迁移学习方法，使模型在面对不同

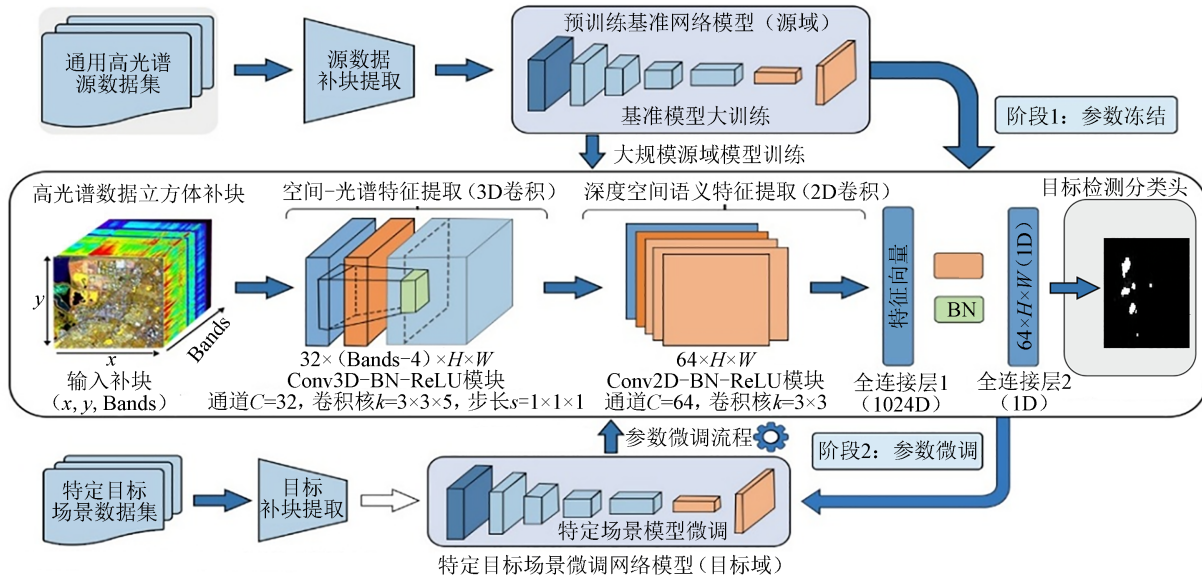


图 3 融合跨域迁移与空谱特征联合提取的 CNN 深度高光谱目标检测网络

Fig. 3 CNN-based deep hyperspectral target detection network integrating cross-domain transfer and spatial-spectral feature joint extraction

时间序列、不确定观测条件时仍具备较强的检测能力。各种方法的优劣如表 2 所示。

表 2 空间-光谱联合深度表征 CNN 方法对比
Table 2 Comparison of CNN methods with spatial-spectral joint deep representation

方法	年份	优势	局限性
3D CNN ^[57]	2019	联合空谱特征且召回率高	计算成本高, 高度依赖先验条件
TCRN ^[58]	2022	独立表征且类内特征紧凑	依赖高分辨率, 对超参数敏感
S2ADet ^[59]	2023	可实现场景级完整边界框定位	降维易产生失真, 对极小目标识别能力有限
3D-2D CNN-TL ^[60]	2024	对背景扰动鲁棒且环境适应性强	极端背景下易出现误报且实时性欠佳

3.1.3 深度结构化推断模型

图 4 展示了基于 CNN 的深度结构化推断模型, 其架构体现了 HTD 中 CNN 的一种演化形式。与作为独立分类模块不同, CNN 被整合到更广泛的建模框架之中, 融合了统计正则化、弱监督学习、多实例学习、迁移学习和模型驱动展开等方法的框架中。其中, CNN 主要提供非线性特征表示能力, 而整体检测性能则源于 3 个组件的系统性融合: 深度特征表示、任务相关的先验知识以及结构化的邻域约束。

由图 4 可见, 结构化 CNN 方法通常包含 3 个关键环节: 首先, 通过样本重组、域对齐或物理建模引入任务先验; 其次, 由 CNN 提取非线性空

间-光谱特征; 最后, 通过注意力、多实例聚合、CEM 类约束或稀疏优化展开生成检测响应。相比普通 CNN, 该类方法的优势在于可解释性更强, 也更适合弱标签、跨传感器和复杂背景条件, 但其代价是模型结构更加复杂, 模型训练稳定性和超参数敏感性更高。

2019 年, Wang 等^[61]率先将 HTD 建模为一个多实例学习问题, 通过一维 CNN 提取谱特征, 并结合自注意力机制选出重要实例, 从而实现了仅在区域级标签下的深度端到端检测, 验证了 CNN 在标签不准确的情况下仍然有效。为解决样本稀缺和模型对传感器存在高度依赖的问题, Shi 等^[62]提出 SDAFL 跨传感器少样本域适配框架, 通过对抗式 CNN 学习相似性嵌入, 结合多维卷积适应空间结构, 将模型检测能力从单域适用推向跨设备可迁移场景。2022 年, Yang 等^[63]提出 DCEM 框架, 将经典的约束能量最小化的物理约束直接融入到 CNN 的目标函数, 采用 Pixel-based 和 Cube-based 模式来验证非线性响应学习的有效性, 证明了 CNN 可以作为一个有统计性质的深度检测算子。2023 年, Jiao 等^[63]提出一种基于 L1 稀疏正则化的多示例注意力网络, 使目标实例的注意力更加凸显, 提高了弱监督检测的理论约束力。为克服样本不平衡问题, Gao 等^[64]提出利用 TBRM 随机波段掩码方法来构造训练样本, 并在 CNN 训练过程中加入残差连接和 SE 通道注意力, 采用三元组损失实现了单先验光谱特征的稳定提取; 此外, Wang 等^[65]利用 AAE 编码器和 CNN 架构增强了奇

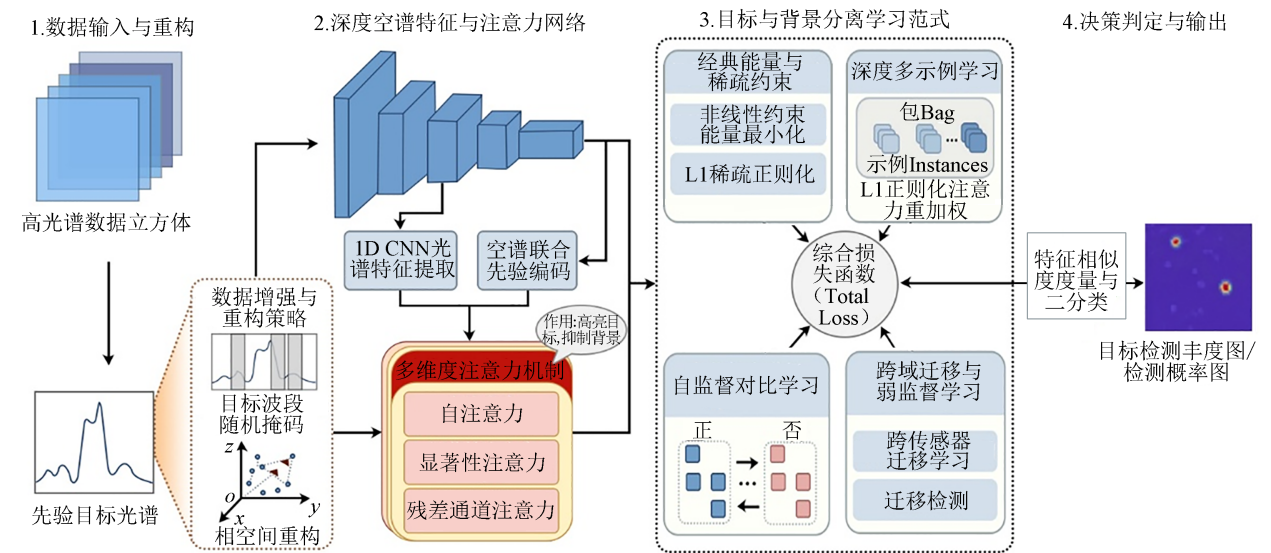


图 4 基于 CNN 的深度结构化推断模型

Fig. 4 A CNN-based deep structured inference model

偶波段采样, 摆脱了样本构造的人工依赖。2023 年, Li 等^[66]提出 TUTDF 框架, 将水体辐射传输机理和域适配 CNN 子网相结合, 实现了从合成数据到真实水下环境的物理迁移。进入新的研究阶段, 模型的可解释性和结构先验成为热点方向。2024 年, Dong 等^[27]提出一种全新的空谱联合稀疏先验编码网络 JSPEN, 通过稀疏重构优化算法逐层展开为 BSPLSn 和 TSPLSn 卷积子网, 使每一个网络层模块都具有明确的物理意义, 弥补了传统 CNN 难以解释的缺陷。2025 年, Imani 等^[67]将光谱序列当作动态系统信号, 用相空间重构法将信号展开后, 交给多分支 CNN 架构处理。这一系列的工作共同表明, 深度高光谱检测中的 CNN 已从最初的末端分类器转变为集非线性映射、约束编码以及响应塑形于一体的结构化核心算子。上述各方法的优缺点如表 3 所示。

3.2 Transformer

3.2.1 度量学习与全监督 Transformer

从传统的统计先验向数据驱动范式转变, 是近年来 HTD 领域的研究热点。然而, 真实视场中目标图谱极度匮乏, 使常规分类网络无法突破这一障碍。为此, 已有研究提出了一种创新性的范式, 即将目标检测任务重新定义为全监督表征和深度度量学习的光谱相似性匹配问题, 其关键核心为: 1) 深度度量学习很大程度上依靠大量的正负样本对来进行特征空间校准; 2) 由于存在只有一条或极少几条先验目标光谱的情况, 该范式采用数据增强技术来完成全监督表征; 3) 底层逻辑通常脱胎于线性混合模型 (linear mixing model,

LMM) 或物理辐射传输模型。图 5 为基于度量学习与全监督 Transformer 的 HTD 原理图。

表 3 基于深度结构化推断模型 CNN 方法
Table 3 CNN methods based on deep structured inference models

方法	年份	优势	局限性
MIL-CNN ^[61]	2019	多实例学习结合自注意力, 用于弱监督检测	仅使用一维卷积, 未利用空间上下文信息
SDAFL ^[62]	2021	半监督域适应结合小样本学习, 支持跨传感器检测	迁移性能严重依赖源域数据的丰富性和质量
DCEM ^[26]	2022	与 CEM 模型结合, 可实现稳健的非线性检测	性能高度依赖网络结构和超参数调节
MIL-CNN-L1 ^[63]	2023	L1 稀疏正则化可有效滤除误检实例	偏向一维光谱特征, 复杂背景下表征能力有限
TBRM ^[64]	2023	目标波段随机遮蔽可有效缓解类别不平衡	遮蔽增强样本可能违背实际物理规律
SCLHTD ^[65]	2023	自监督对比学习消除对高质量标签的依赖	奇偶波段采样可能破坏原始光谱连续性
TUTDF ^[66]	2023	支持从合成数据到真实水下场景的跨域迁移	合成数据与真实数据之间存在分布差异
JSPEN ^[27]	2024	将物理优化过程嵌入网络, 可解释性强	模型迭代展开会增加时间和计算复杂度
PSR-CNN ^[67]	2025	结合相空间重构和显著性图提取非线性特征	相空间重构参数估计带来额外计算开销

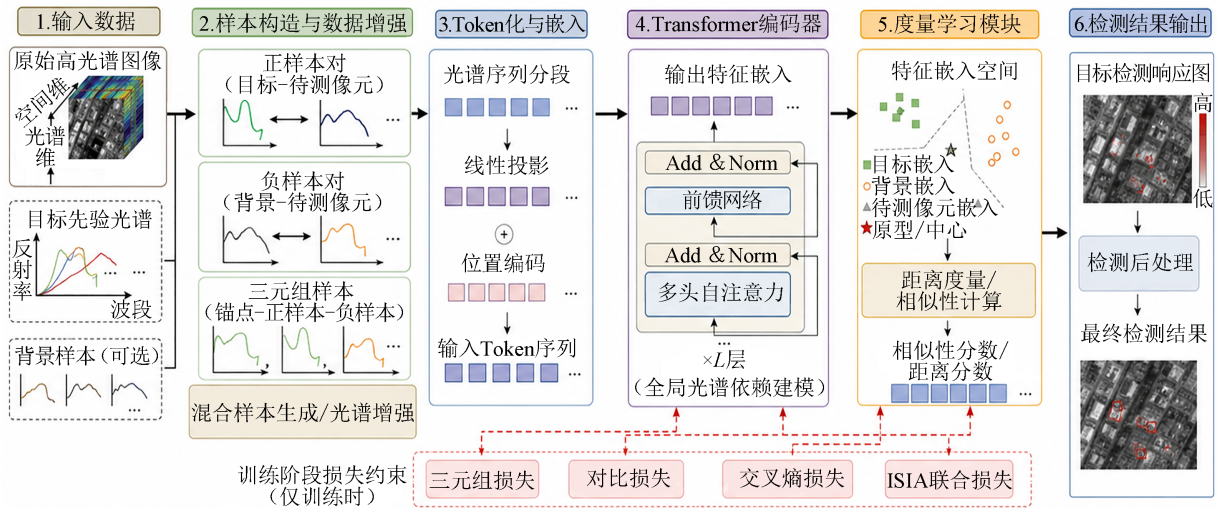


图 5 基于度量学习与全监督 Transformer 的 HTD 原理

Fig. 5 Principle of HTD based on metric learning and fully-supervised Transformer

结合图 5 可知, 该方法的主要贡献在于将 HTD 从“单点统计响应”转化为“全局光谱表征空间中的距离度量”。Transformer 的自注意力机制能够捕获远距离波段间的相关性, 而度量学习机制则缓解了模型在目标样本不足情况下的判别边界学习问题。其局限在于: 全监督样本构造通常依赖目标先验和背景样本纯度, 合成样本与真实混合像元之间可能存在分布差异; 同时, Transformer 计算开销较高, 对训练策略和样本构造质量较为敏感。

2022 年, Rao 等^[68]首次将 Transformer 架构引入到 HTD 领域中, 提出孪生 Transformer 目标检测网络 STTD, 旨在解决传统统计模型过度依赖先验假设以及高光谱真实目标样本稀缺的问题, 通过光谱混合数据增强与孪生网络的结合, 成功提取了光谱的长程依赖特征, 并实现了目标与背景的初步度量分离。2023 年, Jiao 等^[69]提出三元组波段级 Transformer 网络 TSTTD, 克服了早期孪生网络只考虑成对关系且缺乏相邻波段局部特征提取能力等缺陷, 显著提升了模型在复杂背景下的检测精度。2024 年, Yang 等^[70]提出压缩域三元组 Transformer, 解决了直接使用深度学习处理高光谱数据时重建过程耗时较长、特征容易失真的问题。进入 2025 年之后, Liu 等^[71]提出三元组混合注意力网络 THANTD, 针对传统线性插值数据不能真实体现自然界光照、大气变化等问题, 采用物理辐射传输模型来指导模拟样本生成, 获得了更加接近真实物理畸变规律的数据, 同时通过通道和空间混合注意力机制进一步加强了特征表征能力。同年, Cui 等^[72]提出多尺度轻量级 Trans-

former 网络 MSLTD, 通过多尺度稀疏子空间设计, 在降低高光谱信息冗余的同时大幅减小了计算开销。此外, Zheng 等^[73]提出混合 Transformer-Mamba 网络 HTMNet, 首次把具有线性计算复杂度的状态空间模型应用到 HTD 领域, 通过并行的多尺度全局特征提取 Transformer 模块和局部特征提取 Mamba 模块实现了光谱特征多视角的完美交互融合。为打破传统线性数据增强在产生样本多样性和数据真实性等方面的禁锢, Li 等^[74]提出一种利用扩散模型和卷积门控机制的检测算法 HTD-DMCG, 采用前沿的生成式 AI 技术做样本繁育, 并设计了新的目标聚合、背景分离损失函数, 使目标特征的类内紧凑度达到了新高度。上述方法的优缺点如表 4 所示。

3.2.2 弱监督与跨域小样本范式 Transformer

由于具有“图谱合一”的独特优势, 高光谱图像在精细地物识别方面有着巨大的应用潜力。但是, 因光照变化、大气散射等造成的“同物异谱”和“异物同谱”现象, 加之像素级标注成本过高, 导致高度依赖大量样本的传统全监督学习模型难以达到最优效果。近年来, 具有全局感受野的 Transformer 架构被引入 HTD 领域, 并且和弱监督学习以及跨域小样本学习范式紧密结合, 为解决高光谱目标检测所遭遇的“维数灾难”和“样本缺乏”两大难题提供了新的途径。

由图 6 可以看出, 此类方法的核心不是单纯的使用 Transformer, 而是将其作为跨域表征和弱标签推断的中枢。弱监督学习解决像素级标注难的问题, 跨域小样本学习解决不同传感器、不同场景之间的分布偏移问题。其优势在于更贴近真实

应用条件, 但局限也更加突出: 弱标签可能引入噪声, 源域与目标域差异过大时迁移会失效, 且 Token 构造和注意力计算会增加模型复杂度。

表 4 度量学习与全监督 Transformer 方法

Table 4 Metric learning and fully-supervised Transformer methods			
方法	年份	优势	局限性
STTD ^[68]	2022	孪生 Transformer 可捕获全局特征	处理多波段图像时计算成本高
TSTTD ^[69]	2023	光谱级特征提取, 并引入 ISIA 损失	对背景样本提取质量要求高
CS-TTD ^[70]	2024	结合压缩感知, 无需重建图像	压缩率过低时检测性能下降
THANTD ^[71]	2025	结合混合注意力和双重损失	样本构造复杂, 并依赖粗检测结果
MSLTD ^[72]	2025	轻量化多尺度设计, 且计算效率高	对学习超参数较为敏感
HTMNet ^[73]	2025	引入 Mamba 提取全局与局部特征	双分支结构复杂, 且计算和存储需求大
HTD-DMCG ^[74]	2025	扩散模型增强样本, 且可分性高	训练和生成样本需要额外时间

近年来, 许多学者对 Transformer 在 HTD 领域中的应用进行了深入研究, 在解决数据规模、空间特征利用以及跨域泛化等主要问题上取得诸多成果。早在 2022 年, Qin 等^[75]证明了标准视觉大模型在遥感解译上的迁移能力, 提出的 HTD-VIT 方法无需复杂的定制化模块, 即可利用原生视觉 Transformer 强大的表征能力克服不同传感器之间的异质性障碍, 有效地提取出高鉴别力的图谱联合特征。随后, 为解决海量高光谱数据导致的计算雪崩、隐私割裂等问题, Li 等^[76]于 2023 年

创新性地 将分布式深度学习引擎嵌入到 Transformer 中, 提出 HTDFormer 架构, 采用环形全归约拓扑搭建起多节点合作网络, 不仅显著提升了弱监督情形下大量数据的吞吐速度, 更通过梯度聚合有效规避了单个视角下的过拟合问题。在此基础上, Qin 等^[77]针对纯光谱解析容易导致空间上下文断裂的缺点, 在 2024 年提出了双轨并行的 HTD-TS3 网络, 把 Transformer 机制改造成深层次的光谱-空间相似度度量中枢, 成功在极稀疏先验条件下实现了长程空间关联和光谱异质特征的紧密融合, 大幅增强了模型在复杂背景下的抗干扰检出能力。针对跨域场景下的“特征漂移”问题, Feng 等^[78]在 2025 年提出了基于 Transformer 的跨域小样本学习标杆框架 TCFSL, 在特征层和分布层上同时嵌入了域自适应插件, 并通过交叉注意力机制将领域特异性噪声剥离出来, 获得了具有高泛化性的领域不变特征, 为极端缺乏先验知识的新场景目标快检提供了极具应用前景的范例。各种方法的优缺点如表 5 所示。

3.2.3 无监督与生成式自监督范式 Transformer

在 HTD 领域中一直存在缺乏高质量像素级标注数据的固有缺陷。为冲破这一数据壁垒, 学术界将目光投向了无监督学习以及生成式自监督学习 (generative self-supervised learning, GSSL) 范式 Transformer。这类方法的主要思想就是摆脱对人工标注数据的依赖, 从大量的无标注数据中提取出高光谱数据所特有的空间-光谱流形结构。GSSL 通常采用掩码图像建模策略, 通过随机遮蔽一部分数据块, 让网络去重建, 从而使模型内化全局光谱的连续性特征; 而无监督对比学习则是通过在特征空间里博弈, 最大可能地提高同源数

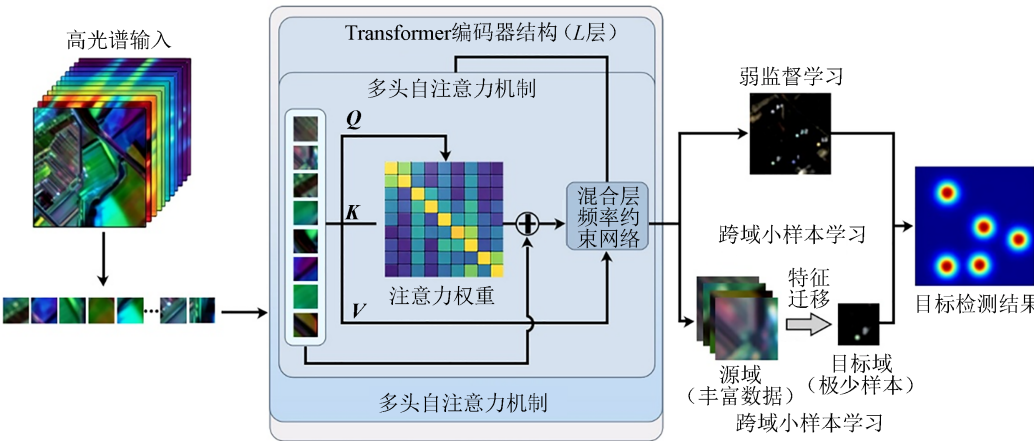


图 6 基于 Transformer 的弱监督与跨域小样本高光谱目标检测原理

Fig. 6 Principle of Transformer-based weakly-supervised and cross-domain few-shot hyperspectral target detection

据之间的相似度，同时减小异源样本之间的特征距离，进而得到可有效对抗复杂背景和“同物异谱”的强判别性表征。基于无监督与生成式自监督范式 Transformer 的 HTD 原理如图 7 所示。

表 5 弱监督与跨域小样本范式 Transformer 方法
Table 5 Weakly-supervised and cross-domain few-shot Transformer methods

方法	年份	优势	局限性
HTD-ViT ^[75]	2022	引入视觉 Transformer 捕获联合空谱特征，并学习每个像素及其邻域的判别特征	依赖粗检测样本的选择精度
HTDFormer ^[76]	2023	采用分布式训练支持大规模数据，并提出基于 Transformer、分布式学习和 Ring All-Reduce 架构的弱监督 HTD 框架	性能受伪标签生成质量限制
HTD-TS3 ^[77]	2024	跨区域建模长程空间依赖，并首次将 Transformer 机制引入 HTD，用于建模长程光谱依赖并进行联合空谱推理	序列长度增加时构造复杂度上升
TCFSL ^[78]	2025	支持跨域小样本学习增强泛化，并通过 Cross-Transformer 在源域和目标域构建 FSL 任务，进行特征级与分布级域适应	严重依赖源域数据的分布质量

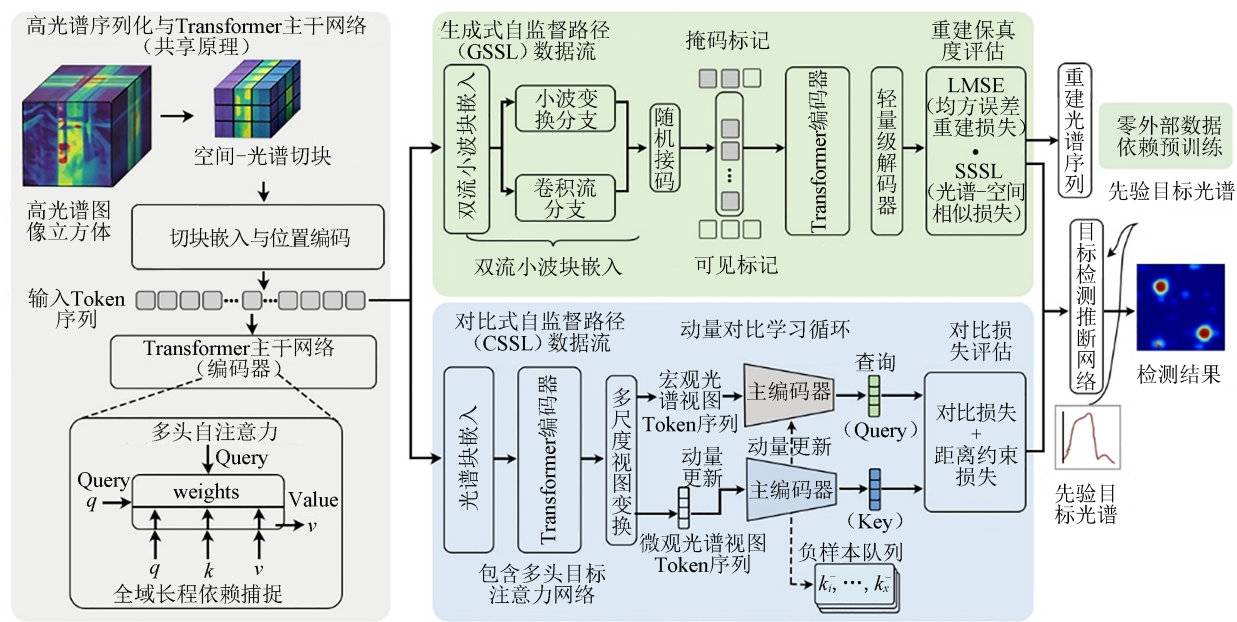


图 7 基于无监督与生成式自监督范式 Transformer 的 HTD 原理

Fig. 7 Principle of HTD based on unsupervised and generative self-supervised Transformer

从图 7 可以看出，无监督和生成式自监督 Transformer 的关键在于“预训练任务是否真的服务于检测”。掩码重建能够强化模型对光谱连续性和空间结构的理解，对比学习能够增强目标与背景在特征空间中的可分性，目标先验约束则有助于避免模型只学习背景重构而忽略弱目标。该类方法适合标注极少或无标注场景，但其风险在于：重建任务可能过于关注背景主成分；负样本选择会影响对比学习效果；生成式预训练的计算成本也较高。

近年来，将 Transformer 的强学习能力同无监督或者自监督范式深度耦合在高光谱目标检测方面取得了重大进展。为解决传统的 Transformer 在全局特征提取时易忽略局部微观光谱细节的问题，

Wang 等^[79]在 2024 年首次提出了基于无监督动量对比学习的 Transformer 网络 MCLT，在没有人工标注的情况下，将局部微末变异和全局长程依赖融合起来，极大地增强了模型在复杂背景下的抗干扰能力。随后，针对复杂成像条件导致的光谱变异以及特征混淆问题，Tian 等^[80]在 2025 年提出一种宏微观光谱对比学习方法 MMSCL，利用多头多目标注意力网络提取得到具有很强判别性的簇元，使得模型能够在宏观包络趋势和微观吸收特征 2 种不同尺度上突出目标与背景之间的差异。与此同时，针对生成式预训练存在“需要依靠大量的外部高光谱数据进行预训练”的弊端，Qin 等^[81]在 2025 年构建了一种结合小波变换的 GSSL 两阶段框架，采用双流小波块嵌入模块以及包含全局-

局部自注意力机制的非对称 Transformer 自编码器, 实现了仅利用单幅待测图像就能进行高质量特征重建的“零外源数据”目标, 并且还开源了包含数万张图像的超大规模预训练数据集 Hyper-Mix-10K。该框架独创性地把目标先验光谱当作刚性约束深深嵌入到 Transformer 的自注意力演算及特征解码流中, 彻底改变了极少先验条件下高光谱靶标“开箱即用”式盲搜的泛化评判标准^[82]。各方法的优缺点如表 6 所示。

表 6 无监督与生成式自监督范式 Transformer 方法
Table 6 Unsupervised and generative self-supervised Transformer methods

方法	年份	优势	局限性
MCLT ^[79]	2024	融合局部与全局特征, 抗干扰能力强	依赖负样本, 计算开销高
MMSCL ^[80]	2025	采用多尺度对比, 增强特征判别性	结构复杂, 特征融合困难
GSSL-WT ^[81]	2025	无需外部源数据, 能有效内化特征	重构任务耗时, 对掩码敏感
SCTS ^[82]	2025	引入硬物理约束, 在先验极少时仍有效	训练过程繁琐, 依赖先验精度

3.3 自编码器

3.3.1 单纯依靠自编码器的方法

高光谱图像数据量较大且具有高维特性, 在处理时往往面临特征冗余与计算复杂度高的难题。AE 作为一种经典的无监督学习模型, 在数据降维与特征表示等方面具有天然优势。AE 主要由编码器和解码器两部分组成。在目标检测任务中, 通过学习背景像素的潜在分布规律, 能够高精度地

重构背景, 而对未参与训练的异常或目标像素重构误差较大, 以此来实现目标的检测与分离。

结合图 8 可知, AE 的目标检测能力并不来自显式分类边界, 而来自“背景可重构、目标难重构”的差异。因此, AE 尤其适合标签缺乏和先验目标不足的场景。但这一假设并非总是成立的, 如果目标与背景光谱高度相似, 或者模型质量过好导致目标也能被重构, 则模型的残差检测能力就会下降。因此, 后续研究通常利用光谱正则、去噪约束、变分建模或注意力机制来限制模型过度重构目标。

近几年, 基于 AE 的高光谱目标检测技术通过不断改善网络结构以及损失函数, 逐步突破了背景干扰和目标变异等相关技术难题。2020 年, Xie 等^[22]针对传统 AE 在复杂背景下特征表示能力不足的问题, 提出 SRUN 方法, 将正则化约束和背景抑制策略相结合, 显著提升了 AE 对背景信息的精确重构能力, 实现了异常目标与背景的高效分离。为克服光谱变异、局部噪声干扰所导致的模型检测精度降低问题, Li 等^[83]于 2022 年提出了 ULM-MDL 模型, 将高光谱图像的光谱信息和空间信息融合到 AE 中, 在改进解码的基础上提高了网络对于局部细节以及全局上下文的捕捉能力。针对高光谱数据的光谱混叠和多重无关特征干扰严重等问题, Shi 等^[84]提出了 ITAE 方法, 在自编码器结构里引入高级特征转换和注意力机制, 有效剔除了冗余和干扰信息, 使模型聚焦于目标最重要、最能区分的特征。针对复杂背景下目标和背景特征边界模糊的问题, Aala 等^[85]提出了 SVHAE 方

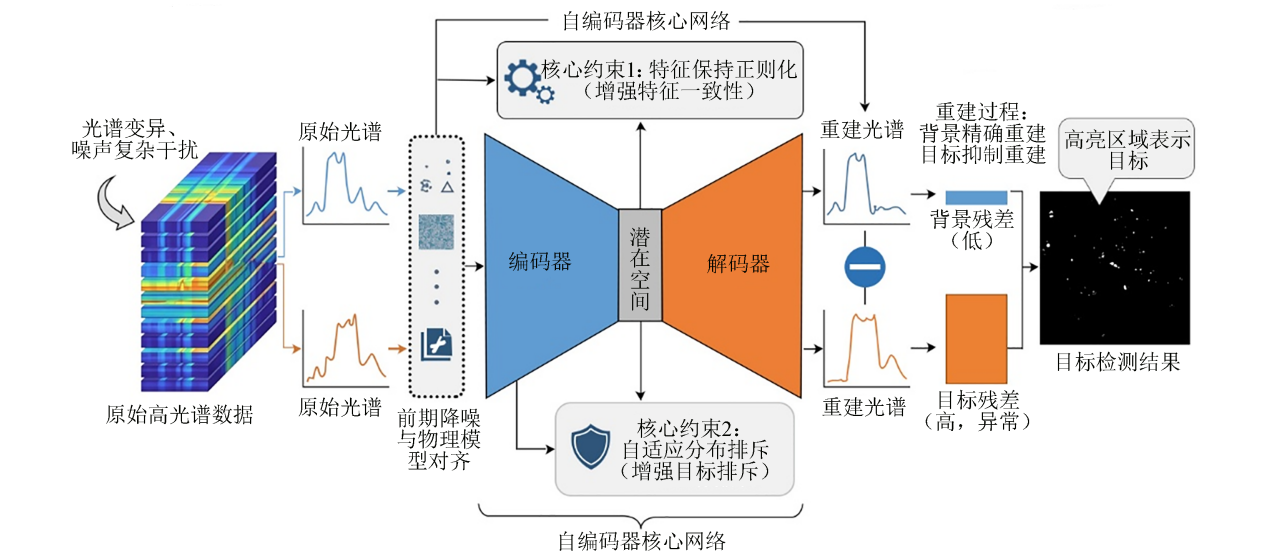


图 8 基于约束自编码器的背景重建与残差检测原理
Fig. 8 Principle of background reconstruction and residual detection based on constrained autoencoder

法，在自编码器的潜在空间上进行更深层次优化，通过细致的分布对齐和特征解耦方法，实现了对目标特征的纯净分离。纵观这一发展脉络，AE 在高光谱目标检测领域中的特征提取能力不断得到加强、重构机制日趋完备，这些革新办法逐步攻克了从背景抑制、空谱融合到特征解耦等核心难题。各方法的优缺点如表 7 所示。

表 7 自编码器方法
Table 7 Autoencoder-based methods

方法	年份	优势	局限性
SRUN ^[22]	2020	引入光谱正则化项（光谱角距离）增强特征光谱一致性	隐层节点判别能力有限，需要额外特征选择
ULMMDL ^[83]	2022	逐层去噪，无需预设分布假设，采用统一空谱端元提取	依赖背景端元提取质量，计算成本高
ITAE ^[84]	2024	基于截断 KL 散度的迭代优化，且参数自适应，无需字典辅助	迭代过程复杂，收敛判定受阈值影响
SVHAE ^[85]	2025	双路径解码（线性+非线性），可有效处理水下光谱变化（吸收/散射）	专为水下场景设计，陆地场景泛化仍需验证

3.3.2 自编码器与其他深度学习技术相结合

HTD 领域面临的主要难题是怎样在纷繁复杂的大背景下，将目标特征与背景光谱分离。不同于传统仅将光谱维度作为输入的方法，现代方法更倾向于在潜在空间里把目标和背景进行解耦。该类方法一般采用 AE 作为核心结构，在编解码的

过程中加入物理约束、空间-光谱联合以及背景先验等手段，使得网络得到的潜在变量由简单的压缩表示转变为具有明确检测导向的特征描述。

从图 9 可以看出，AE 与其他深度学习技术结合后，检测逻辑由“重构异常”扩展为“潜空间解耦”。一方面，空间-光谱模块可以补充邻域结构信息，缓解单像素重构时的不稳定性；另一方面，物理约束和目标先验可以防止网络只学习背景主成分而忽略弱目标。此外，非线性解码、变分约束和注意力机制有助于处理复杂混合和光谱变异。该类方法的优势是可同时兼顾无监督学习、物理解释和非线性建模，但模型复杂度较高，且不同约束项之间的权重设置会显著影响检测结果。

近年来，基于 AE 的高光谱目标检测算法已由单个光谱特征提取发展为时空谱多维融合。Shi 等^[86]在 2019 年提出三维宏观-微观残差自编码器 3D-MMRAE，采用双支路协同建模的方式将空间和光谱联合特征融入到检测体系中，从而克服了仅依靠单一维度光谱判别的不足。同年，Shi 等^[87]提出判别式约束堆叠自编码器 DCSSAE，在传统结构上增加了目标-背景距离约束，并将重构学习和检测判别性相结合，为后来的研究奠定了理论基础。2021 年，针对高光谱数据中的维度冗余问题，Shi 等^[88]又提出了基于 RoI 特征变换和多尺度光谱注意力的深度空间-光谱网络 DSSN，通过兴趣区域重整显著提升了模型在复杂背景下的检测鲁棒性。同年，为解决水下特种环境探测问题，Qi

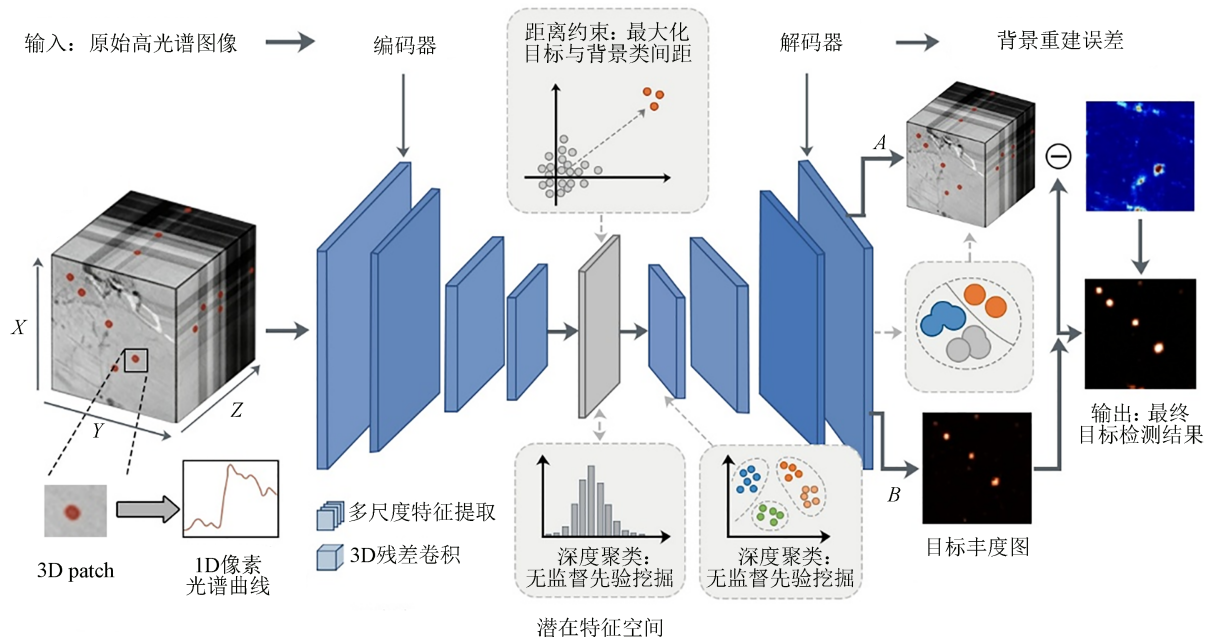


图 9 融合潜空间约束的自编码器目标检测框架

Fig. 9 Autoencoder-based target detection framework integrating latent space constraints

等^[89]提出了UTD-Net, 将深层物理模型和异常分离策略融入到解码过程中, 增强了模型在水下复杂介质环境中的适应性。2023年, 针对目标先验缺失的问题, Zhang等^[23]提出了SSDCTD法, 该方法不依赖外部先验, 而是通过自监督学习来挖掘数据内部的结构规律。2024年, Tian等^[24]提出正交子空间引导变分自编码器OS-VAE, 其主要思想是构造出极其纯净的背景表示, 用背景生成建模的方法反向突出目标信号。同年, Liu等^[25]针对近岸场景的非线性混合问题, 提出NUN-UTD模型, 通过非线性解码机制进一步打破传统的线性混合假设。AE已从单纯的无监督工具转变为一个包含空间约束、注意力机制、物理模拟以及非线性解混的综合性检测框架, 实现了检测精度和环境适应性的双重提升。上述各种方法的优缺点如表8所示。

表 8 自编码器与其他深度学习方法相结合的方法
Table 8 Methods combining autoencoder with other deep learning techniques

方法	年份	优势	局限性
3D-MMRAE ^[86]	2019	融合多尺度空谱特征, 同时可以提取宏观和微观特征	训练极其耗时, 且存在样本不平衡问题
DCSSAE ^[87]	2019	扩大目标-背景距离, 在特征空间中最大化差异	忽略空间相关信息
DSSN ^[88]	2021	抑制冗余, 定位精确, 采用多尺度光谱注意力	亚像素检测性能一般
UTD-Net ^[89]	2021	物理意义强, 基于测深模型消除干扰	高度依赖水体参数精度
SSDCTD ^[23]	2023	无需标注, 通过深度聚类前置任务进行自监督挖掘	聚类超参数难以设置
OS-VAE ^[24]	2024	模型分布刻画准确, 通过正交子空间损失实现强背景抑制	需要粗检测预处理, 并且相对复杂
NUN-UTD ^[25]	2024	提取弱信号, 适应性强, 采用伪混合数据训练策略	深度模型易过拟合

4 实验与分析

4.1 实验条件设置

4.1.1 高光谱图像数据集

圣地亚哥高光谱数据集由位于圣地亚哥的机载可见光和红外成像光谱仪 (airborne visible infrared imaging spectrometer, AVIRIS) 获取。图10展示了San Diego数据集图像。原始图像包含224

个光谱通道, 波长范围为370~2 510 nm。在实验中, 保留了189个波段, 排除了水吸收波段和低信噪比波段。整个城市景观图像大小为400像素×400像素, 空间分辨率为3.5 m。实验中使用的AVIRIS-1和AVIRIS-2分别为200像素×200像素和120像素×120像素的子图像, 如图10(b)所示。检测目标为子图像中的3架飞机。在本实验中, 将这3架飞机的平均光谱作为目标光谱。

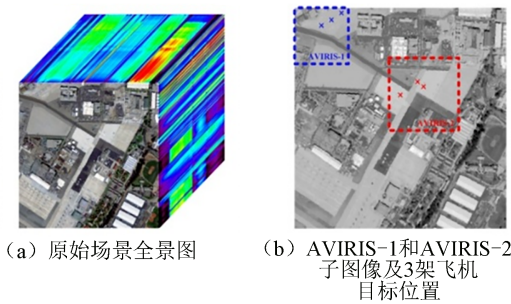


图 10 San Diego 数据集

Fig. 10 San Diego dataset

Cuprite数据图像于1997年在美国内华达州Cuprite矿区获取。该图像中约有14种矿物, 包括布丁石、钠蒙脱石、诺顿石 (铁质粘土)、高岭石-瓦克斯石等。图11展示了多用途扫描仪3.3软件生成的矿物图, 本文使用该图像中250像素×191像素的子集进行实验, 如图中红色框所示。目标为铵长石 (buddingtonite), 约占39个像素。数字光谱库中有2种铵长石的光谱, 本文取其平均值作为所有算法的目标光谱。表9展示了3个数据集的详细数据。

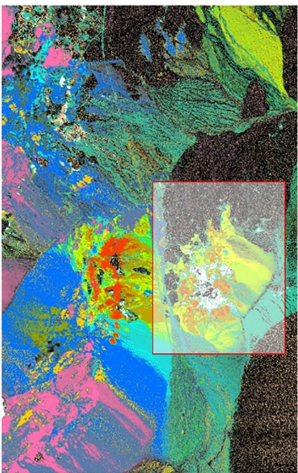


图 11 Cuprite 数据集

Fig. 11 Cuprite dataset

4.1.2 实验环境与参数设置

为确保实验的公平性与可复现性, 统一设置实验环境与核心参数。具体配置如下: 硬件为1台

搭载 Intel Core i7/i9 系列处理器、NVIDIA GeForce RTX 5060 Laptop GPU (8.5 GB 显存) 的笔记本电脑。软件环境的操作系统为 Windows11 22H2 (版本 10.0.26200)，深度学习框架采用 PyTorch 2.9.0 (CUDA 12.8)，配套 cuDNN v9.1.0。所有深度学习模型均采用 Adam 优化器，初始学习率设为 10^{-4} ，Batch Size 为 64，最大训练轮次为 200，并采用早停策略防止过拟合。输入数据均进行逐波段标准化，即减去全局均值并除以标准差。

表 9 实验数据集基本信息

Table 9 Basic information of experimental datasets			
数据集	传感器	波段数	空间尺寸
Sandiego1	AVIRIS	189	200 像素×200 像素
Sandiego2	AVIRIS	189	120 像素×120 像素
Cuprite	AVIRIS	188	250 像素×191 像素

4.2 具有代表性的高光谱图像目标检测算法的对比实验研究

为评估各阶段不同典型算法的性能，依据以下原则筛选了 10 种代表性算法进行实验对比。一

是不同时期的代表性算法：涵盖传统统计方法、早期深度学习方法及当前三大主流深度学习范式；二是不同技术代表性算法：所选算法均为各自技术路线下的开创性或性能标杆算法；三是文献可复现：优先选择已公开源代码的算法，以确保实验的公平性与可验证性。

在本节中，采用 AUC 作为评估指标，以比较不同算法的性能。图 12~14 为不同典型方法在不同数据集下的目标检测可视化结果，其定量评价结果如表 10 所示

由表 10 可得，因传统光谱匹配算法 MIACE 仅基于测量像素光谱和目标光谱的角度差来进行目标检测，未利用空间信息，因此在所有数据集上的 AUC 值均低于 0.7，表明传统方法难以处理具有高维度、强冗余性、混合像素问题的高光谱数据。CEM 类算法 (E-CEM、hCEM) 采用最小化能量约束来优化特征提取，在一定程度上提高了目标和背景之间的区分度，AUC 值在 0.75~0.81 之间，但仍受困于“手工特征泛化能力差”的缺陷，难以适应复杂场景中地面物体分布的差异。卷积神经网络、转换器、自动编码器等 AI 算

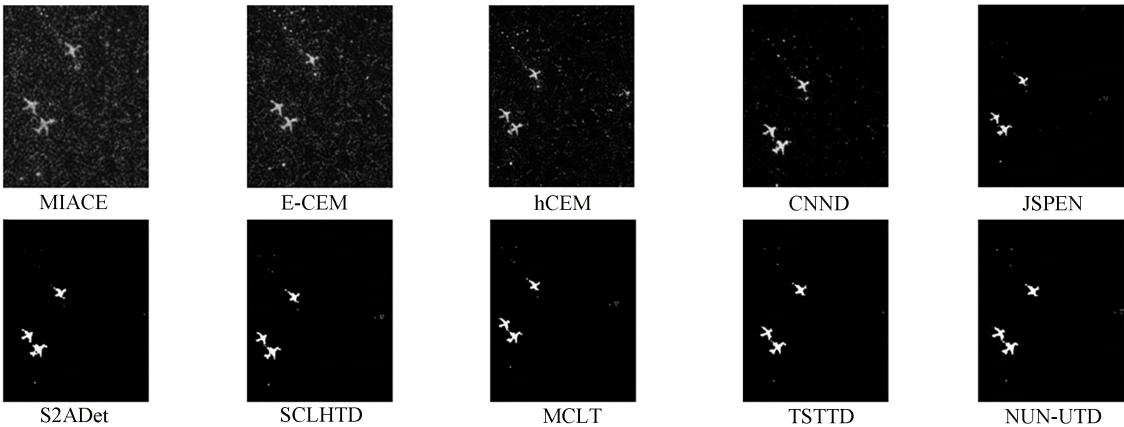


图 12 各方法在 Sandiego1 (AVIRIS-2) 数据集下的实验结果

Fig. 12 Experimental results of various methods on Sandiego1 (AVIRIS-2) dataset

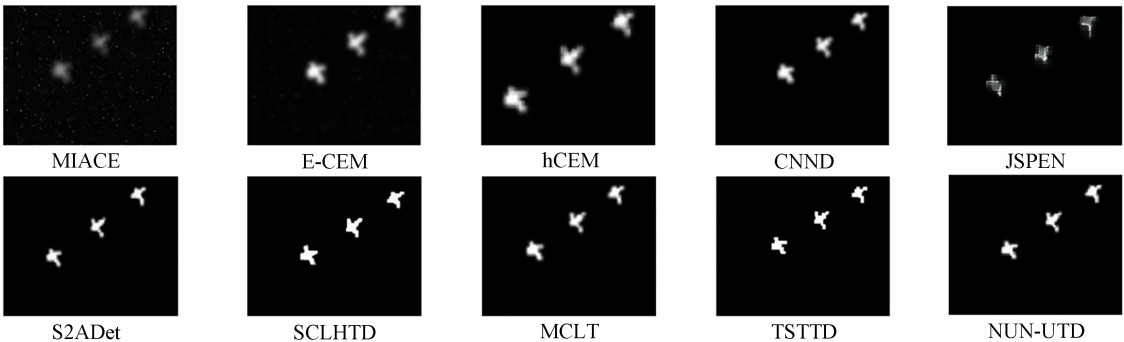


图 13 各方法在 Sandiego2 (AVIRIS-1) 数据集下的实验结果

Fig. 13 Experimental results of various methods on Sandiego2 (AVIRIS-1) dataset

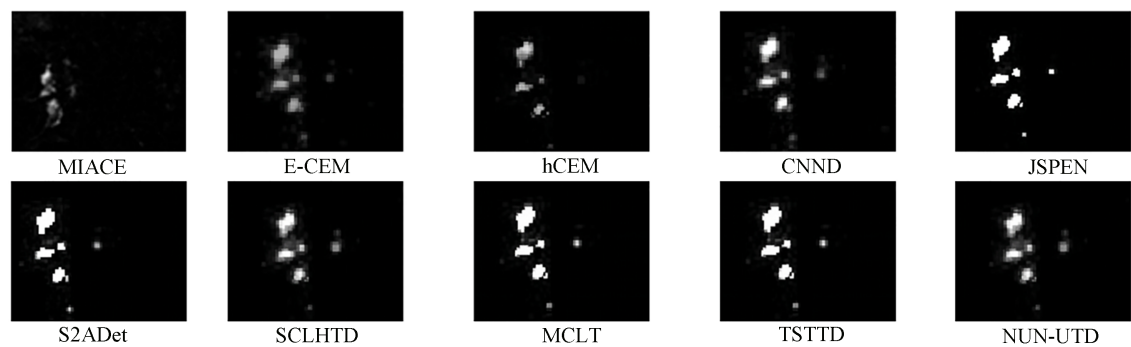


图 14 各方法在 Cuprite 数据集下的实验结果

Fig. 14 Experimental results of various methods on Cuprite dataset

法凭借数据驱动的特征学习能力, AUC 值普遍大于 0.86, 证明了“深度学习和 Transformer 架构给高光谱目标检测带来革命性突破”的观点。

表 10 不同算法在高光谱数据集上的性能比较

Table 10 Performance comparison of different algorithms on hyperspectral datasets

方法	Sandiego1	Sandiego2	Cuprite
MIACE	0.681 9	0.675 2	0.699 5
E-CEM	0.758 3	0.736 9	0.794 3
hCEM	0.775 3	0.792 3	0.805 7
CNND	0.803 7	0.813 2	0.802 4
JSPEN	0.819 9	0.829 1	0.864 3
S2ADet	0.864 2	0.867 2	0.894 5
SCLHTD	0.876 3	0.897 4	0.828 3
MCLT	0.877 2	0.819 9	0.883 1
TSTTD	0.889 2	0.880 5	0.873 8
NUN-UTD	0.895 2	0.861 9	0.831 2

注: 粗体为每列最优值。

值得一提的是, 各种典型的人工智能算法在不同数据集上的表现差异较大, 这也深刻揭示了“算法架构同数据集特征的契合程度才是决定检测性能的关键要素”。以 CNN 架构为基础的 SCL-
HTD 模型, 凭借多头注意力机制, 可以有效捕捉到长距离光谱的相关性以及全局的空间信息, 在背景分散、混合像素比例高的 Sandiego2 数据集上表现最优, AUC 值达 0.897 4。以自编码器为框架的 NUN-UTD 模型在 Sandiego1 数据集上达到最优性能表现, AUC 值高达 0.895 2, 其编解码过程利用无监督重建来去除冗余、净化关键特征, 显著提升了检测的鲁棒性。而 CNN 类的 S2ADet 模型在目标空间分布比较规则、局部光谱特征显著的 Cuprite 数据集中表现最好, AUC 值达 0.894 5,

充分验证了卷积网络结合空间金字塔池化提取局部特征的优势。进一步结合图 12~14 的可视化结果可以得出, 传统光谱匹配算法 MIACE 和 CEM 类算法识别出的目标像素与背景的边界较为模糊, 且在背景中存在大量椒盐噪声。而基于卷积神经网络、转换器、自动编码器等智能目标检测算法的图像质量得到大幅提升。

然而, 本次实验的复现结果整体低于原始文献的报告值。经过溯源分析可知, 该种性能损耗并非源于核心算法代码本身, 而是由数据预处理、运行环境、实验配置 3 个方面存在的非算法逻辑上的偏差所导致的。首先, 复现实验未完全和原始数据预处理相对应, 缺少某种光谱幅值缩放、均值方差调整或者数据集划分策略, 导致输入模型的数据分布出现偏差; 其次, 由于运行环境以及数据库版本的不同, 间接影响了模型的收敛轨迹; 最后, 硬件配置和训练超参数的不同, 使得模型在单 GPU 和默认配置下难以找到全局最优解, 从而导致运行相同代码时性能存在较大差距。

5 研究展望与结论

HTD 的发展过程实际上就是一场从“人工建模”到“自主学习”的范式迁移。未来技术发展将从以下 4 个方向演进。

1) 创新模型架构。HTD 中传统深度学习架构的局限日益凸显, CNN 的固定卷积核难以适应多尺度光谱变化, Transformer 全局注意力机制因维度越来越大而成本过高, 通用型自动编码器由于没有任务导向而缺少特征判别能力。就目前而言, 未来的主流发展方向为自适应混合架构, 即通过引入可变形卷积赋予网络动态感知能力, 采用稀疏注意力机制优化算力分配, 并重构编码器的特征表达。尤其值得关注的是, 图神经网络与

Transformer 的深度融合, 可用以表征波段间的复杂高阶相关性, 为高光谱数据处理开辟了一条极具前景的新途径。

2) 优化深度网络。由于深度网络简单堆叠造成的梯度消失、训练不稳定、特征退化等问题, 单纯的增加层数已经不能满足性能要求。针对 CNN 特征衰减、Transformer 收敛难、AE 易过拟合等现实挑战, 智能深度控制技术成为破解难题的重要手段。未来将采用神经架构搜索精确设定 CNN 深度, 依靠自适应调整来优化 Transformer 拓扑结构, 借助变分推理提升自动编码器的层次表征能力, 并结合知识蒸馏技术将重型模型的表征优势“往下兼容”到轻量化架构上, 为复杂模型在实际使用中的高效部署提供可行方案。

3) 创新特征学习方法。针对目前特征学习中 CNN 偏向局部纹理、Transformer 忽略波段细节、自编码器容易受到主流地物干扰的问题, 未来的发展重点是多粒度特征融合。采用注意力引导机制, 使 CNN 可以精准捕捉判别性波段, 通过谱感知架构增强 Transformer 的光谱细节保持能力, 使用对比学习提高自编码器的特征辨识度, 从根源上改善表征质量。在此基础上, 结合元学习框架赋予模型未知场景下快速迁移能力, 通过多任务学习实现特征共享, 从而扩大特征泛化范围, 构建出更加坚韧的高光谱智能探测系统。

4) 优化训练结构。为解决深度学习对标注数据的过度依赖问题, 尤其是当样本数量过少时, 监督学习失效、自监督算法不稳定等现象, 未来的训练方式会朝着自监督和弱监督的深度融合方向发展。通过对比学习驱动 CNN 挖掘无标签数据中的深层表征、利用掩码谱建模引导 Transformer 学习光谱序列的内在结构以及结合传统自编码器强化特征提取能力, 可以有效解决数据匮乏问题。在此基础上, 采用课程学习逐步构建推进知识体系, 同时借助联邦学习来克服多源数据之间的“孤岛”现象, 将会给创建高效协同的高光谱智能感知系统提供重要的策略支持。

参考文献

- [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 Qingxi, Zhang Bing, Zhang Lifu. Current progress of hyperspectral remote sensing in China[J]. National Remote Sensing Bulletin, 2016, 20(5): 689-707.
- [2] Man Qixia, Dong Pinliang, Guo Huadong. Pixel- and feature-level fusion of hyperspectral and lidar data for urban land-use classification[J]. International Journal of Remote Sensing, 2015, 36(6): 1618-1644.
- [3] Schmid T, Rodríguez-Rastrero M, Escribano P, et al. Characterization of soil erosion indicators using hyperspectral data from a Mediterranean rainfed cultivated region[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(2): 845-860.
- [4] Harb M M, Dell Acqua F. Remote sensing in multirisk assessment: improving disaster preparedness[J]. IEEE Geoscience and Remote Sensing Magazine, 2017, 5(1): 53-65.
- [5] Harsanyi J C, Chang Chein-I. Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach[J]. IEEE Transactions on Geoscience and Remote Sensing, 1994, 32(4): 779-785.
- [6] Harsanyi J C. Detection and classification of sub-pixel spectral signatures in hyperspectral image sequences[D]. Baltimore, MD, USA: University of Maryland, Baltimore County, 1993: 127-128.
- [7] Ren Hsuan, Du Qian, Chang Chein-I, et al. Comparison between constrained energy minimization based approaches for hyperspectral imagery[C]// IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data. Piscataway, NJ, USA: IEEE, 2003: 244-248.
- [8] Zhao Rui, Shi Zhenwei, Zou Zhengxia, et al. Ensemble-based cascaded constrained energy minimization for hyperspectral target detection[J]. Remote Sensing, 2019, 11(11): 1310.
- [9] Robey F C, Fuhrmann D R, Kelly E J, et al. A CFAR adaptive matched filter detector[J]. IEEE Transactions on Aerospace and Electronic Systems, 1992, 28(1): 208-216.
- [10] Kraut S, Scharf L L. The CFAR adaptive subspace detector is a scale-invariant GLRT[J]. IEEE Transactions on Signal Processing, 1999, 47(9): 2538-2541.
- [11] Jiao Xiuli, Chang Chein-I. Kernel-based constrained energy minimization (K-CEM)[J]. Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV, 2008, 6966: 69661S.
- [12] Kwon H, Nasrabadi N M. Kernel orthogonal subspace projection for hyperspectral signal classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(12): 2952-2962.
- [13] Kwon H, Nasrabadi N M. Kernel adaptive sub-

- space detector for hyperspectral imagery[J]. *IEEE Geoscience and Remote Sensing Letters*, 2006, 3(2): 271-275.
- [14] Chen Yi, Nasrabadi N M, Tran T D. Sparse representation for target detection in hyperspectral imagery[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2011, 5(3): 629-640.
- [15] Li Wei, Du Qian, Zhang Bing. Combined sparse and collaborative representation for hyperspectral target detection[J]. *Pattern Recognition*, 2015, 48(12): 3904-3916.
- [16] Zhang Lefei, Zhang Liangpei, Tao Dacheng, et al. Hyperspectral remote sensing image subpixel target detection based on supervised metric learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(8): 4955-4965.
- [17] Li Wei, Wu Guodong, Du Qian. Transferred deep learning for hyperspectral target detection[C]//2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Piscataway, NJ, USA: IEEE, 2017: 5177-5180.
- [18] Du Jinming, Li Zhiyong, Sun Hao. CNN-based target detection in hyperspectral imagery[C]//2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Piscataway, NJ, USA: IEEE, 2018: 2761-2764.
- [19] Zhu Dehui, Du Bo, Zhang Liangpei. Two-stream convolutional networks for hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(8): 6907-6921.
- [20] Xie Weiying, Zhang Jiaqing, Lei Jie, et al. Self-spectral learning with GAN based spectral-spatial target detection for hyperspectral image[J]. *Neural Networks*, 2021, 142: 375-387.
- [21] Xu Shufang, Geng Sijie, Xu Pengfei, et al. Cognitive fusion of graph neural network and convolutional neural network for enhanced hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5515915.
- [22] Xie Weiying, Yang Jian, Lei Jie, et al. SRUN: spectral regularized unsupervised networks for hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(2): 1463-1474.
- [23] Zhang Xiaodian, Gao Kun, Wang Junwei, et al. Self-supervised learning with deep clustering for target detection in hyperspectral images with insufficient spectral variation prior[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2023, 122: 103405.
- [24] Tian Qu, He Chengxun, Xu Yang, et al. Hyperspectral target detection: learning faithful background representations via orthogonal subspace-guided variational autoencoder[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5516714.
- [25] Liu Jiaxuan, Qi Jiahao, Zhu Dehui, et al. Detecting nearshore underwater targets with hyperspectral nonlinear unmixing autoencoder[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5529615.
- [26] Yang Xiaoli, Zhao Min, Shi Shuaikai, et al. Deep constrained energy minimization for hyperspectral target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 8049-8063.
- [27] Dong Wenqian, Wu Xiaoyang, Qu Jiahui, et al. Deep spatial: spectral joint-sparse prior encoding network for hyperspectral target detection[J]. *IEEE Transactions on Cybernetics*, 2024, 54(12): 7780-7792.
- [28] Keshava N, Mustard J F. Spectral unmixing[J]. *IEEE Signal Processing Magazine*, 2002, 19(1): 44-57.
- [29] Chang Chein-I, Heinz D C. Constrained subpixel target detection for remotely sensed imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2000, 38(3): 1144-1159.
- [30] Heinz D C, Chang Chein-I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2001, 39(3): 529-545.
- [31] Yang Shang, Shi Zhenwei. SparseCEM and Sparse-eACE for hyperspectral image target detection[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(12): 2135-2139.
- [32] Zou Zhengxia, Shi Zhenwei. Hierarchical suppression method for hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(1): 330-342.
- [33] Shang Xiaodi, Song Meiping, Wang Yulei, et al. Target-constrained interference-minimized band selection for hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(7): 6044-6064.
- [34] Zhang Jing, Zhao Rui, Shi Zhenwei, et al. Bayesian constrained energy minimization for hyperspectral target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 8359-8372.

- [35] Zhao Xiaobin, Hou Zengfu, Wu Xia, et al. Hyperspectral target detection based on transform domain adaptive constrained energy minimization[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 103: 102461.
- [36] Fu Tongtong, Fu Xiaoning. An improved CEM target detection algorithm for hyperspectral images[J]. Applied Physics, 2019, 9(2): 63-70.
- [37] Zhang Yuxiang, Du Bo, Zhang Liangpei. Regularization framework for target detection in hyperspectral imagery[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(1): 313-317.
- [38] Shang Xiaodi, Song Meiping, Wang Yulei, et al. Hyperspectral target detection based on target-constrained interference-minimized band selection[C]//2020 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ, USA: IEEE, 2020: 2033-2036.
- [39] Sun Xudong, Zhang Hongjun, Xu Fengqiang, et al. Constrained-target band selection with subspace partition for hyperspectral target detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 9147-9161.
- [40] Chen Jie, Chang Chein-I. Background-annihilated target-constrained interference-minimized filter (TCIMF) for hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5540224.
- [41] Tian Zhiyong, Gao Kun, Zhang Xiaodian. Constrained-target band selection based on band combination for hyperspectral target detection using CEM [C]//2022 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ, USA: IEEE, 2022: 771-774.
- [42] Wang Teng, Du Bo, Zhang Liangpei. An automatic robust iteratively reweighted unstructured detector for hyperspectral imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(6): 2367-2382.
- [43] Zare A, Jiao Changzhe, Glenn T. Discriminative multiple instance hyperspectral target characterization [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(10): 2342-2354.
- [44] Wang Teng, Du Bo, Zhang Liangpei. A kernel-based target-constrained interference-minimized filter for hyperspectral sub-pixel target detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2013, 6(2): 626-637.
- [45] Liu Xiaofang, Yang Chun. A kernel spectral angle mapper algorithm for remote sensing image classification[C]//2013 6th International Congress on Image and Signal Processing (CISP). Piscataway, NJ, USA: IEEE, 2013: 814-818.
- [46] Du Bo, Zhang Yuxiang, Zhang Liangpei, et al. Beyond the sparsity-based target detector: a hybrid sparsity and statistics-based detector for hyperspectral images[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5345-5357.
- [47] Chen Yi, Nasrabadi N M, Tran T D. Kernel sparse representation for hyperspectral target detection[C]//2012 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ, USA: IEEE, 2012: 7484-7487.
- [48] Shi Yanzi, Li Jiaojiao, Li Yunsong, et al. Hyperspectral target detection using a bilinear sparse binary hypothesis model[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5519113.
- [49] Zhang Yuxiang, Du Bo, Zhang Liangpei. A sparse representation-based binary hypothesis model for target detection in hyperspectral images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(3): 1346-1354.
- [50] Zhang Yuxiang, Zhang Liangpei, Du Bo, et al. A nonlinear sparse representation-based binary hypothesis model for hyperspectral target detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(6): 2513-2522.
- [51] Ling Qiang, Guo Yulan, Lin Zaiping, et al. A constrained sparse-representation-based binary hypothesis model for target detection in hyperspectral imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(6): 1933-1947.
- [52] Zhao Xiaobin, Li Wei, Zhang Mengmeng, et al. Adaptive iterated shrinkage thresholding-based Lp-norm sparse representation for hyperspectral imagery target detection[J]. Remote Sensing, 2020, 12(23): 3991.
- [53] Zeng Jiang, Wang Qunming. Sparse tensor model-based spectral angle detector for hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5539315.
- [54] Du Jinming, Li Zhiyong. A hyperspectral target detection framework with subtraction pixel pair features[J]. IEEE Access, 2018, 6: 45562-45577.
- [55] Zhang Gaigai, Zhao Shizhi, Li Wei, et al. HTD-Net: a deep convolutional neural network for target detection in hyperspectral imagery[J]. Remote

- Sensing, 2020, 12(9): 1489.
- [56] Feng Zhenyuan, Zhang Junping, Feng Jian. Spectral-spatial joint target detection of hyperspectral image based on transfer learning[C]//2020 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ, USA: IEEE, 2020: 1770-1773.
 - [57] Freitas S, Silva H, Almeida J M, et al. Convolutional neural network target detection in hyperspectral imaging for maritime surveillance[J]. International Journal of Advanced Robotic Systems, 2019, 16(3): 1729881419842991.
 - [58] Li Yanshan, Chen Shifu, Xu Jianjie, et al. A spatial-spectral feature based target detection framework for high-resolution HSI[J]. International Journal of Remote Sensing, 2022, 43(15/16): 5775-5799.
 - [59] He Xiao, Tang Chang, Liu Xinwang, et al. Object detection in hyperspectral image via unified spectral-spatial feature aggregation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5521213.
 - [60] Zhao Jiale, Wang Ganglong, Zhou Bing, et al. Exploring an application-oriented land-based hyperspectral target detection framework based on 3D-2D CNN and transfer learning[J]. EURASIP Journal on Advances in Signal Processing, 2024, 2024(1): 37.
 - [61] Wang Xiuxiu, Chen Xiaoying, Gou Shuiping, et al. Hyperspectral target detection via deep multiple instance self-attention neural network[C]//2019 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ, USA: IEEE, 2019: 2284-2287.
 - [62] Shi Yanzi, Li Jiaojiao, Li Yunsong, et al. Sensor-independent hyperspectral target detection with semisupervised domain adaptive few-shot learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(8): 6894-6906.
 - [63] Jiao Changzhe, Chen Chao, Gou Shuiping, et al. L1 sparsity-regularized attention multiple-instance network for hyperspectral target detection[J]. IEEE Transactions on Cybernetics, 2023, 53(1): 124-137.
 - [64] Gao Hongmin, Zhang Yitong, Chen Zhonghao, et al. Hyperspectral target detection via spectral aggregation and separation network with target band random mask[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5515516.
 - [65] Wang Yulei, Chen Xi, Zhao Enyu, et al. Selfsupervised spectral-level contrastive learning for hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5510515.
 - [66] Li Zheyong, Li Jinghua, Zhang Pei, et al. A transfer-based framework for underwater target detection from hyperspectral imagery[J]. Remote Sensing, 2023, 15(4): 1023.
 - [67] Imani M, Cerra D. Phase space deep neural network with saliency-based attention for hyperspectral target detection[J]. Advances in Space Research, 2025, 75(4): 3565-3588.
 - [68] Rao Weiqiang, Gao Lianru, Qu Ying, et al. Siamese transformer network for hyperspectral image target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5526419.
 - [69] Jiao Yinyue, Gong Zhiqiang, Zhong Ping. Triplet spectralwise transformer network for hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5519817.
 - [70] Yang Qingyao, Wang Xiaoqin, Chen Long, et al. CS-TTD: triplet transformer for compressive hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5533115.
 - [71] Liu Ju, Wang Haoqian, Cheng Xiangai, et al. THANTD: triplet hybrid attention network for hyperspectral target detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 8831-8844.
 - [72] Cui Ying, Chen Tao, Wang Liguang, et al. Multi-scale lightweight transformer network with adaptive contrastive learning for hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5529416.
 - [73] Zheng Xiaosong, Kuang Yin, Huo Yu, et al. HT-MNet: hybrid transformer-mamba network for hyperspectral target detection[J]. Remote Sensing, 2025, 17(17): 3015.
 - [74] Li Zhaokui, Mo Fan, Zhao Xiaobin, et al. Hyperspectral target detection using diffusion model and convolutional gated linear unit[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5512313.
 - [75] Qin Haonan, Xie Weiying, Li Yunsong, et al. HTD-ViT: spectral-spatial joint hyperspectral target detection with vision transformer[C]//2022 IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ, USA: IEEE, 2022: 555-558.
 - [76] Li Yunsong, Qin Haonan, Xie Weiying. HTD-

- Former: hyperspectral target detection based on transformer with distributed learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5524715.
- [77] Qin Haonan, Xie Weiying, Li Yunsong, et al. HTD-TS3: weakly supervised hyperspectral target detection based on transformer via spectral-spatial similarity[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35 (11): 16816-16830.
- [78] Feng Shou, Wang Xueqing, Feng Rui, et al. Transformer-based cross-domain few-shot learning for hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2025, 63: 1-16.
- [79] Wang Yulei, Chen Xi, Zhao Enyu, et al. An unsupervised momentum contrastive learning based transformer network for hyperspectral target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 17: 9053-9068.
- [80] Tian Jiacheng, Shen Dunbin, Kong Wenfeng, et al. Hyperspectral target detection based on macro-micro spectrum contrastive learning[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2025, 18: 15157-15169.
- [81] Qin Han, Wang Shuo, Li Yunsong, et al. Hyperspectral target detection based on generative self-supervised learning with wavelet transform[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2025, 63: 5510815.
- [82] Qin Haonan, Wang Shuai, Li Yunsong, et al. A signature-constrained two-stage framework for hyperspectral target detection based on generative self-supervised learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2025, 63: 5514917.
- [83] Li Yunsong, Shi Yanzi, Wang Keyan, et al. Target detection with unconstrained linear mixture model and hierarchical denoising autoencoder in hyperspectral imagery[J]. *IEEE Transactions on Image Processing*, 2022, 31: 1418-1432.
- [84] Shi Yidan, Li Xiaorun, Chen Shuhan, et al. Target-driven iterative autoencoder for hyperspectral target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5505518.
- [85] Aala S, Kumar Sikkakolli S, Chinnadurai S, et al. SVHAE: spectral variability-aware hybrid autoencoder for hyperspectral underwater target detection[J]. *IEEE Access*, 2025, 13: 129161-129171.
- [86] Shi Yanzi, Li Jiaojiao, Yin Yaping, et al. Hyperspectral target detection with macro-micro feature extracted by 3-D residual autoencoder[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12 (12): 4907-4919.
- [87] Shi Yanzi, Lei Jie, Yin Yaping, et al. Discriminative feature learning with distance constrained stacked sparse autoencoder for hyperspectral target detection[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(9): 1462-1466.
- [88] Shi Yanzi, Li Jiaojiao, Zheng Yuxuan, et al. Hyperspectral target detection with RoI feature transformation and multiscale spectral attention[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(6): 5071-5084.
- [89] Qi Jiahao, Gong Zhiqiang, Xue Wei, et al. An unmixing-based network for underwater target detection from hyperspectral imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 5470-5487.

Hyperspectral Target Detection:The Evolution from Conventional to Artificial Intelligence

ZHANG Zhili, SUN Haixiang*, DING Yao, LU Zhenhang,
LI Huaichun, LI Jian, WU Junhao

(Rocket Force University of Engineering, Xi'an 710025, China)

ABSTRACT: To systematically analyze the development trajectory of hyperspectral target detection (HTD) technology from traditional mathematical modeling to deep learning frameworks, the performance of various representative algorithms was quantitatively evaluated. The study performed a horizontal compari-

son of classic models such as constrained energy minimization and adaptive matched filtering, as well as comprehensively analyzed the core innovations of three cutting-edge intelligent mechanisms: reshaping a convolutional neural network (CNN) into nonlinear relationship measurers based on joint spatial-spectral representation; precisely capturing long-range spectral dependencies across the entire image by combining the multi-head self-attention mechanism of the Transformer with a generative self-supervised paradigm; and achieving deep decoupling of weak targets from complex backgrounds through multi-dimensional physical constraints in the latent space of autoencoders. Based on the Sandiego1, Sandiego2, and Cuprite datasets, a detailed validation of 10 different algorithms using the area under the curve (AUC) as the core evaluation metric was conducted. The results indicate that, constrained by single-pixel feature fitting and high-dimensional redundancy, traditional methods generally struggle to exceed an AUC of 0.81; in contrast, deep learning models achieve a comprehensive AUC exceeding 0.86. The CNN-based SCLHTD model achieves a peak AUC of 0.8974 on the mixed-background Sandiego2 dataset due to its outstanding long-range correlation capture ability; the CNN-based S2ADet algorithm achieves an AUC of 0.8945 on the spatially regular target-distributed Cuprite dataset; and the unsupervised reconstruction-integrated auto-encoder model NUN-UTD also achieves the best AUC of 0.8952 on the Sandiego1 dataset. As single network models have reached the bottleneck of feature representation, the future breakthrough in hyperspectral detection technology lies in constructing adaptive hybrid architectures and multi-granularity feature fusion mechanisms, with a shift toward unsupervised and weakly supervised paradigms to address the problems of complex environment generalization and few-shot feature transfer.

KEYWORDS: hyperspectral image; target detection; deep learning; image processing

(编辑: 刘建民)