

KDTGAN: 基于 Transformer-GAN 和知识蒸馏的高光谱目标检测

谢雯, 闪晨超, 张哲哲, 张嘉鹏
(西安邮电大学通信与信息工程学院 西安 710121)

摘要: 高光谱目标检测在地球观测中至关重要, 被广泛应用于军事和民用领域。然而, 由于高光谱图像的背景复杂性和目标样本的有限性, 该任务面临较大的挑战。本文首先采用 CEM (约束能量最小化) 粗检测方法提取背景数据。随之, 引入了一种新的知识蒸馏模型, 即 KDTGAN (通过 Transformer-GAN 实现)。教师模型的生成器采用了 Transformer 编码器的结构, 并结合多尺度数据融合的方法, 能够准确地学习背景分布, 进而通过重构背景信息实现目标检测。为了克服 GAN (生成对抗网络) 训练不稳定的挑战, 特别是纯背景数据的稀缺性, 本文提出了一种新的损失算法, 以减小可疑目标样本对模型性能的负面影响。为了降低模型的计算负担, 本文引入知识蒸馏, 并设计新的蒸馏损失对学生模型加以约束, 使模型轻量化的同时提高学生模型检测精度。实验结果表明: KDTGAN 相较于当前检测方法表现更优, 具有更高的检测精度和鲁棒性。

关键词: 高光谱图像; 目标检测; 知识蒸馏; 生成对抗网络; Transformer-GAN

中图分类号: TP751; TP18 **文献标志码:** A **文章编号:** 2095-1000(2024)02-0010-08

DOI: 10.12347/j.ycyk.20240119001

引用格式: 谢雯, 闪晨超, 张哲哲, 等. KDTGAN: 基于 Transformer-GAN 和知识蒸馏的高光谱目标检测[J]. 遥测遥控, 2024, 45(2): 10-17.

KDTGAN: Knowledge Distillation via Transformer GAN for Hyperspectral Target Detection

XIE Wen, SHAN Chenchao, ZHANG Zhezhe, ZHANG Jiapeng

(School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

Abstract: Hyperspectral target detection is crucial in Earth observation for both military and civilian applications. However, complex backgrounds and the scarcity of target samples pose challenges in hyperspectral image analysis. In this paper, we first employ the CEM coarse detection method to extract background data. Subsequently, a novel knowledge distillation model, namely KDTGAN (implemented through Transformer-GAN), is introduced. The generator of this teacher model adopts the structure of a Transformer encoder and combines it with a multi-scale data fusion approach to accurately learn the background distribution, which in turn enables target detection by reconstructing the background information. To overcome the challenge of unstable GAN training, especially the scarcity of pure background data, we propose a new loss algorithm to reduce the negative impact of suspicious target samples on model performance. To reduce the computational burden of the model, we introduce knowledge distillation and design a new distillation loss to constrain the student model to lighten the model while improving the student model's detection accuracy. The experimental results show that KDTGAN performs better than current detection methods with higher detection accuracy and robustness.

Keywords: Hyperspectral; Target detection; Knowledge distillation; GAN; Transformer-GAN

Citation: XIE Wen, SHAN Chenchao, ZHANG Zhezhe, et al. KDTGAN: Knowledge Distillation via Transformer GAN for Hyperspectral Target Detection[J]. Journal of Telemetry, Tracking and Command, 2024, 45(2): 10-17.

基金项目: 国家自然科学基金 (61901365, 62071379); 陕西省自然科学基金 (2019JQ-377); 陕西省教育厅专项科研计划 (19JK0805); 西安邮电大学西邮新星团队项目 (xyt2016-01)

收稿日期: 2024-01-19; **修回日期:** 2024-03-05

0 引言

高光谱图像呈现为一个三维立方体, 蕴含了光谱和空间信息。与RGB图像不同, 高光谱图像的波段范围可达400 nm~2 500 nm^[1], 因其丰富的光谱信息, 在多个应用领域发挥着重要作用。目标检测作为一项基本任务, 在民用和军事等领域受到广泛关注。传统高光谱目标检测方法通常依赖于目标与背景的光谱差异。例如, 约束能量最小化(Constrained Energy Minimization, CEM)^[2]通过施加目标和背景的约束, 构建有限脉冲响应滤波器, 使得滤波器输出能量最小化。基于稀疏和协同表示(Combined Sparse and Collaborative Representation, CSCR)^[3]的组合方法则利用目标和背景的表达残差生成检测输出。自适应相关估计(Adaptive Correlation Estimation, ACE)^[4]和匹配滤波器(Matched Filter, MF)^[5]将目标检测视为假设检验问题。然而, 传统方法往往忽略了高光谱图像所覆盖的大尺度地面场景的复杂性, 其中包含具有不同光谱特性的各种物质^[6]。

随着深度学习方法的发展, 出现了许多基于生成对抗网络(Generative Adversarial Networks, GAN)^[7]、自编码器(Autoencoder, AE)、变分自编码器(Variational Autoencoder, VAE)^[8]和对抗自编码器(Adversarial Autoencoder, AAE)的目标检测方法^[9-13]。这些方法通过利用深度学习的重构特性, 对高光谱数据进行了有效的数据增强, 尤其在无标签高光谱数据集上表现显著。例如, 秦等人提出的PTGAN模型^[14], 其通过设计双阶段加权GAN模型, 充分利用GAN模型的重构能力, 并且借用指数非线性融合策略, 结合了两个阶段的鉴别器特征来抑制背景实现高光谱目标检测任务。然而, 先前的方法^[15, 16]主要基于CNN(Convolutional Neural Networks, 卷积神经网络)捕获高光谱图像的特征, 而CNN的局部感受野机制使其难以学习远程的空谱信息。因此, 一些学者在高光谱目标检测中引入了Transformer模型^[17-20], 以关注全局信息。例如, RAO等人提出的STTD模型^[21], 采用孪生网络的方法, 其两条支路采用相同的Transformer结构, 支路之间权重共享, 最后通过相似度打分得到最终检测结果。LI等人提出的HTD-Former模型^[22], 采用分布式学习的方法, 通过多进程支路并行的方式, 减少了Transformer庞大的计算量问题。

但是, 这些仅使用光谱信息的模型在有效区分地面目标方面存在不足。鉴于此, 本文提出了KDTGAN模型, 主要工作包括:

① 采用CEM粗检测方法选择背景样本进行训练, 在此基础上, 挖掘Transformer在高光谱目标检测任务中背景学习的潜力。为此, 提出了一个基于Transformer和GAN的模型框架, 并通过整合跨尺度数据, 充分考虑了数据在多个尺度上的变化, 提高了模型的适应性, 以有效地利用背景信息和全局高光谱信息。

② 为了解决训练过程目标样本造成的负面影响, 提出了目标抑制损失函数, 使生成结果更趋近于纯背景样本。

③ 采用知识蒸馏的方法对模型进行轻量化处理, 提出了高效的蒸馏损失函数, 通过将学生模型引入来替代教师模型, 从而在减少模型计算量的同时提升学生模型性能。

本文首先, 详细介绍了提出的KDTGAN模型, 其次, 给出了实验结果和相应分析, 最后是本文结论。

1 提出的方法

本文提出了一种通过背景学习的GAN来进行目标检测。所提出的方法体现在三个主要方面: 首先, 提取背景训练样本, 基于KDTGAN对网络进行训练, 以获得唯一的背景表示; 然后, 通过加入蒸馏损失, 对学生模型加以约束, 降低模型计算量同时提升学生模型重构性能; 最后, 用训练好的模型来重建背景, 通过光谱角距离度量来实现目标检测。

2 模型结构

如图1所示, 提出的KDTGAN模型包括: 教师模型生成器(图1(a))、学生模型生成器(图1(b))和教师模型鉴别器(图1(c))。

2.1 样本选择

由于目标训练样本相对于背景训练样本较少, 因此采用以CEM算法为基础的粗检测方法提取背景训练样本。高光谱数据集 $\mathbf{X} \in R^{H \times W \times C}$ 可以用高度 H , 宽度 W , 光谱带数为 C 的三维立方体表示。通过设定阈值将其划分为目标样本 $\mathbf{X}_T \in R^{H_t \times W_t \times C}$ 和背景样本 $\mathbf{X}_B \in R^{H_b \times W_b \times C}$ 。然后, 从背景样本中提取训练样本和测试样本。

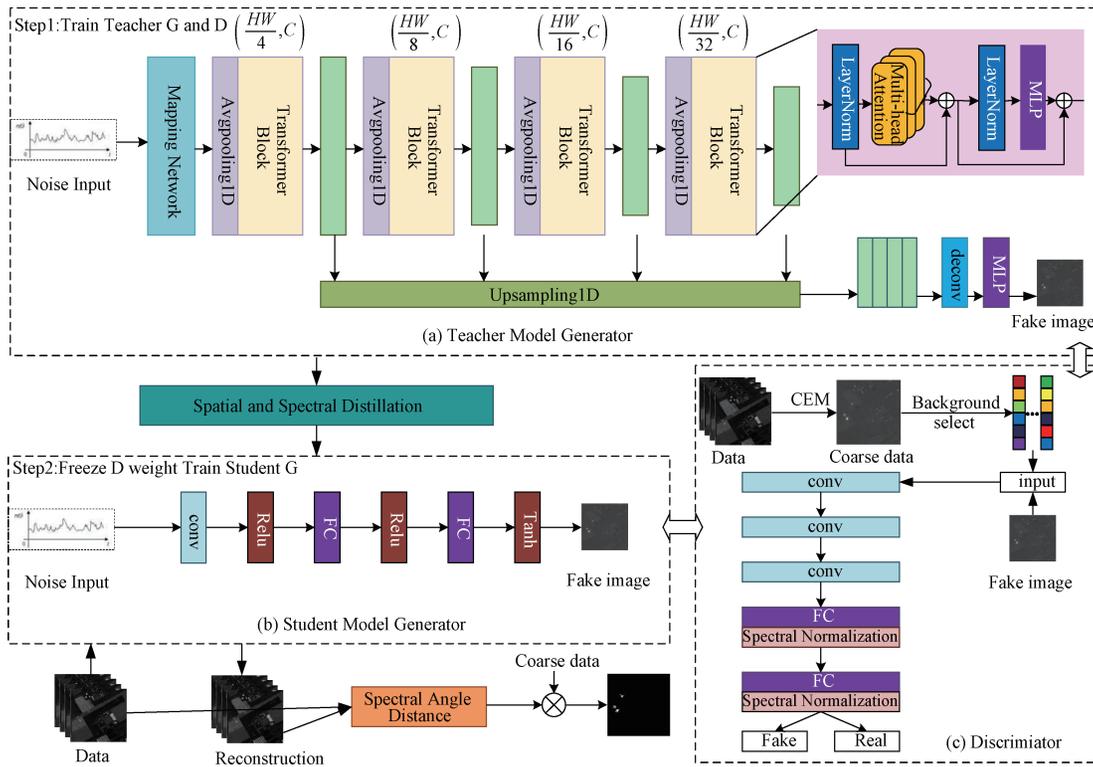


图 1 KDTGAN 模型总体框架图

Fig. 1 Framework diagram of our proposed model

2.2 教师模型

2.2.1 教师模型生成器

教师模型采用了 GAN 模型结构, 其中包括生成器 $G_T(\cdot)$ 和鉴别器 $D_T(\cdot)$, 它们通过对抗博弈方式实现数据生成任务。设计的教师模型生成器的流程图如图 1(a) 所示。首先, 引入了 Mapping Network 层^[23], 其中包含 8 层全连接层, 以挖掘数据的深度特征信息; 接着, 对输出结果进行序列化展平处理, 并分别通过 [4, 8, 16, 32] 尺寸在通道维度的均值池化, 实现多尺度数据划分; 最后, 使用 Transformer 结构对不同尺度的数据进行特征提取, 并采用数据融合模块, 将中间层信息进行加权融合。本文的 Transformer 模型结构参考了文献 [24] 中的模型设计, 多头自注意力机制 (MHSA) 可以描述为 query、key、value 点乘映射, 其中 query、key、value 和最后输出都是向量形式, 输出计算为 value 的加权和, 而权重则由 query 和 key 之间的相应函数计算而来。

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Soft max} \left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} \right) \mathbf{V} \quad (1)$$

式 (1) 中的 $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 分别代表的 query、key、

value 矩阵, d_k 是输入数据的维度。具有 m 个头数的多头自注意力表示为:

$$\text{MHSA}(\mathbf{X}) = \text{Concat}(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_m) \mathbf{W} \quad (2)$$

其中, $\mathbf{h}_i = \text{Attention}(\mathbf{X}\mathbf{W}_i^Q, \mathbf{X}\mathbf{W}_i^K, \mathbf{X}\mathbf{W}_i^V)$, \mathbf{W} 代表相应的可学习矩阵。将每阶段 Transformer 输出的结果 $\mathbf{X}_z = \left\{ (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i) | \mathbf{x}_i \in \mathbf{R}^{\frac{HW}{2^z} \times C} \right\}$, 分别沿空间维度 1D 上采样为统一大小, 再以通道维度拼接和 1D 反卷积实现信息融合, 最后通过 MLP 层, 以得到生成器生成的伪数据 $\hat{\mathbf{Y}}_1$ 。

2.2.2 教师模型鉴别器

教师模型鉴别器结构如图 1(c) 所示, 由三层卷积层和两层全连接层组成, 卷积采用 1D 卷积, 卷积核设置为 3。为了稳定鉴别器的训练过程, 在鉴别器中加入了光谱归一化层, 以稳定鉴别器的权重, 避免出现梯度消失或梯度爆炸的问题。对于鉴别器的输入, 采用经过 CEM 粗检测筛选的背景训练样本进行特征学习, 通过 GAN 的交互来促使生成器学到与背景分布相关的特征。

首先, 训练鉴别器网络, 使其能够区分真实样本和生成器生成的伪数据, 为伪数据样本设定伪标签信息, 通过计算交叉熵损失, 优化鉴别器

的参数，以增强其对生成器输出的对抗性。同时，训练生成器网络，将背景数据作为目标，使生成器能够学到背景分布的特征。生成器的输入是从潜在空间中随机采样的噪声。通过优化生成器的参数，使得生成器能够生成与背景数据相似的合成数据，这一训练过程的目标在于通过对抗训练提高生成器生成数据的逼真度，同时使其能够捕捉背景数据的特征。

2.2.3 目标抑制损失

由于模型训练目的是学习背景样本分布，而提取背景训练样本过程是采用CEM算法进行粗略检测得到，所以训练样本中仍可能是会存在目标样本进而影响背景分布的学习，造成负面影响。为了克服此问题，本文提出了一个损失函数 L_{STT} ，与传统的损失函数相比，该方法在处理高光谱图像时能够更有效捕捉复杂的空间特征。通过度量生成器输出的伪数据 $\widehat{\mathbf{Y}}_i$ 和目标先验 $\mathbf{D}_i = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_i\} \in \mathbf{R}^{1 \times C}$ 之间的相似性，抑制可能存在的目标样本，损失函数 L_{STT} 如下：

$$L_{\text{STT}} = \frac{1}{K} \sum_{i=1}^K \text{Relu} \left(\frac{\sum_{i=1}^N (\widehat{\mathbf{y}}_i - \bar{\mathbf{y}}_i)(\mathbf{d}_i - \bar{\mathbf{y}}_i)}{\|\widehat{\mathbf{y}}_i - \bar{\mathbf{y}}_i\|_2 \|\mathbf{d}_i - \bar{\mathbf{y}}_i\|_2} \right) \quad (3)$$

$\widehat{\mathbf{y}}_i$ 表示生成器得到的伪结果， $\bar{\mathbf{y}}_i$ 表示对 $\widehat{\mathbf{y}}_i$ 取均值， \mathbf{d}_i 表示目标先验样本。通过余弦相似度度量二者的相似度比值，即计算出生成器生成的伪数据样本是否和真实的目标样本接近，产生由目标样本干扰而得到的结果。其中皆要减去均值 $\bar{\mathbf{y}}_i$ ，目的是使得向量更关心方向而不是幅度，加上Relu激活函数，剔除小于0的值，仅保留正值，并且通过Top函数，取前 K 个向量的均值，以得到损失值进而抑制目标样本的存在。

2.3 学生模型

由于教师模型采用了Transformer结构，导致计算量较大，但独立使用学生模型结构的检测结果却表现较差。因此，为了在降低模型计算量的同时不损失模型的性能，采用了知识蒸馏的方法。在这个过程中，以离线蒸馏的方式协助学生网络生成器 $G_s(\cdot)$ 与已训练好的 $D_T(\cdot)$ 进行博弈，仅更新生成器权重而不对鉴别器权重进行更新。学生网络生成器如图1(b)所示，采用一层卷积层、两层全连接层和非线性激活函数构建。然而，蒸馏方式的选择会极大程度影响蒸馏效果，无监督模型需

要从数据中学到更通用和抽象的表示，而这些表示通常存在于网络的中间层。因此，引用关系知识蒸馏(Relational Knowledge Distillation, RKD)的思想^[25]，输出 $D_T(\cdot)$ 的 $(M-1)$ 个中间层特征和最终结果，并分别按照空间维度和光谱维度将知识传递给学生网络，如图2所示。同样，对学生网络的生成器输出 $(M-1)$ 个中间层特征和最终结果，将两者都输入损失函数，以实现蒸馏过程。因此，所设计的蒸馏损失如下：

$$D(x, y) = \sqrt{(x-y)^T S^{-1} (x-y)} \quad (4)$$

$$\varphi_d = \frac{1}{N} \left(D(\widehat{\mathbf{y}}_i^t, \widehat{\mathbf{y}}_i^s) - D(\widehat{\mathbf{y}}_i^t, \mathbf{d}_i) + \theta, 0 \right) \quad (5)$$

$$L_d = \sum_i^M (\varphi_d^1, \varphi_d^2, \dots, \varphi_d^M) \quad (6)$$

公式(4)为马氏距离，其中 S 表示输入数据的协方差矩阵。公式(5)中， $\widehat{\mathbf{y}}_i^t, \widehat{\mathbf{y}}_i^s$ 分别表示教师网络和学生网络输出， \mathbf{d}_i 为目标先验， θ 为设定的固定值，使得二者距离保持固定距离间隔，防止距离过小而梯度消失， $(\cdot)_+$ 代表了差值小于0时，以0取代，不算入损失。其中，将二者做差计算，确保同一类别之间的间距小于不同类别的距离，实现同类接近、异类远离。公式(6)表示将 $G_T(\cdot)$ 输出的 $(M-1)$ 个中间层特征和最终结果计算的损失值进行求和，以得到最终蒸馏损失。

高光谱数据集为三维立方体结构，包含空间和光谱信息，因此在知识蒸馏过程中，将教师网络的知识传递给学生时通过空间和光谱两个维度分别进行，如图2所示。通过3D卷积设置不同的卷积核，提取到不同的特征信息，其中空间维度卷积核的 C 表示通道维数，再经过批归一化和Relu非线性激活函数稳定数据分布，最后残差连接以提取到不同的特征信息。

由于学生生成器和教师网络鉴别器进行对抗性训练，因此在优化学生模型时加入经过Sigmoid的交叉熵损失，最终学生模型的损失函数表示为，

$$L_s = L_{\text{BCE}} + L_d \quad (7)$$

在训练阶段，生成器接收随机噪声作为输入，这有助于引入一定程度的随机性和多样性，从而使生成器能够生成更具创造性的样本。鉴别器接收原始背景分布的数据，以区分生成器生成的样本与真实样本的差异。在最后检测阶段，由于生成器已经在训练中学到了背景分布的特征，将高光谱数据集 \mathbf{X} 输入到已学习到背景分布的 $G_s(\cdot)$ 中，

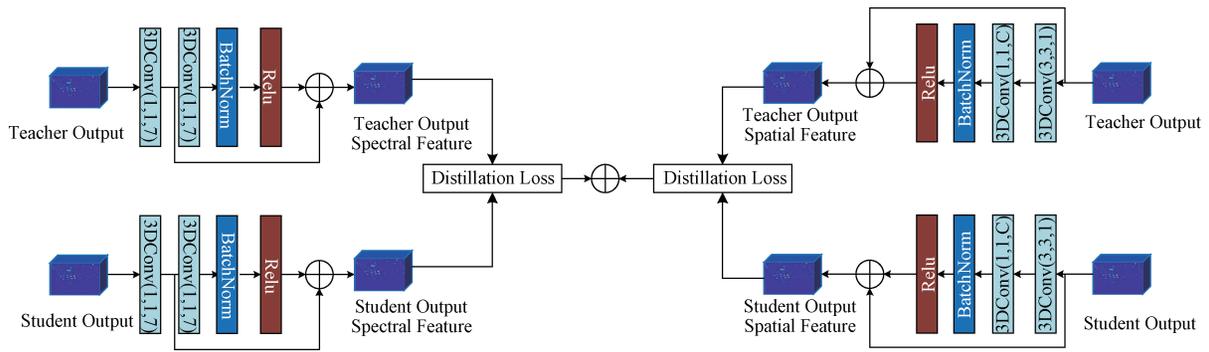


图2 空间信息和光谱信息知识蒸馏

Fig. 2 Distillation of spatial and spectral information knowledge

得到的重构结果可以被视为生成器尝试去除目标样本, 重构出符合训练中学到的背景分布的样本, 实现对目标和背景样本的重构, 重构的结果反映了生成器在学习阶段所学到的背景分布特征, 同时将目标样本视为“噪声”(在生成器学到的背景分布之外的信息), 以产生与训练数据相符的生成结果, 扩大目标样本和背景样本分布的差异性, 得到重构结果与高光谱数据集 X 通过光谱角距离度量, 获得权重值, 最终再与粗检测结果进行点乘以得到检测结果图。

3 实验

3.1 实验设置

所有实验均在配备 NVIDIA RTX 4070Ti GPU 和 Intel Core i9-13900K CPU 的台式机上进行, 并在 San Diego、EI Segundo 和 Synthetic 高光谱数据集上测试。采用自适应时刻估计法(Adaptive Moment Estimation, Adam)以对模型权重优化, 其中 epoch 设置 500, 生成器学习率设定为 0.000 1, 鉴别器的学习率为 0.000 01。使用曲线下的面积(Area Under the Curve, AUC)值和接收机工作特征曲线(Receiver Operating Characteristic curve, ROC)来显示检测精度和效果。

3.2 数据集

San Diego: 第一个数据集是由 AVIRIS 传感器获得的, 覆盖了美国圣地亚哥机场范围, 空间大小为 200 像素×200 像素, 包括 224 个光谱通道, 波长从 370 nm~2 510 nm。空间分辨率约为 3.5 m, 去除吸水带和低信噪比波段后, 共使用了 189 个波段, 三架飞机被视为目标像素占 134 个像素值。

EI Segundo: 第二个数据集是由 AVIRIS 传感器捕获, 它有 224 个光谱通道, 波长从 366 nm~

2 496 nm。这个城市场景覆盖了美国加州的一个地区。空间大小为 250 像素×300 像素, 其中主要几个居民区, 公园和炼油区被视为目标像素, 占 2 048 像素。

Synthetic: 第三个数据集是由美国地质调查局(USGS)的数字光谱库合成的高光谱数据集, 空间大小为 64 像素×64 像素, 其中包含有 15 种矿物端元的数据合成, 选择 Labradorite 作为目标, 像素值占 12。

3.3 结果分析

3.2.1 对比实验

为了评估提出模型的检测性能, 本文采用 CEM、ACE、MF、SID、SAM 和 BLTSC 检测方法进行比较, 如图 3 所示。其中, 图 3(a)、图 3(b)分别为 San Diego、EI Segundo 和 Synthetic 数据集的伪彩图和地面真值图, 图 3(c)~图 3(j)依次为 SID、SAM、CEM、ACE、MF、BLTSC、KDTGANS-、KDTGANS+ 在数据集上的检测图。

特别注意的是, 本文在 Synthetic 数据集测试时加入了信噪比为 20 dB 的高斯白噪声, 使用 KDTGAN 表示教师模型, KDTGANS+ 表示蒸馏后的学生模型, KDTGANS- 表示未蒸馏的学生模型, 可以观察到我们提出的 KDTGANS+ 模型突出了目标, 与其他检测方法相比对背景有抑制作用。为了对检测性能进行详细的定性分析, 本文绘制了 (P_D, P_F) 的对数尺度 ROC 曲线, 如图 4 所示, 以用来评估检测和背景抑制性能, 并将检测结果进行了可视化操作, 对应的 AUC 值如表 1 所示。从图 4 可以明显看出, KDTGANS+ 模型具有较好的背景抑制能力。

$$SNR(\text{dB}) = 10 \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right) \quad (8)$$

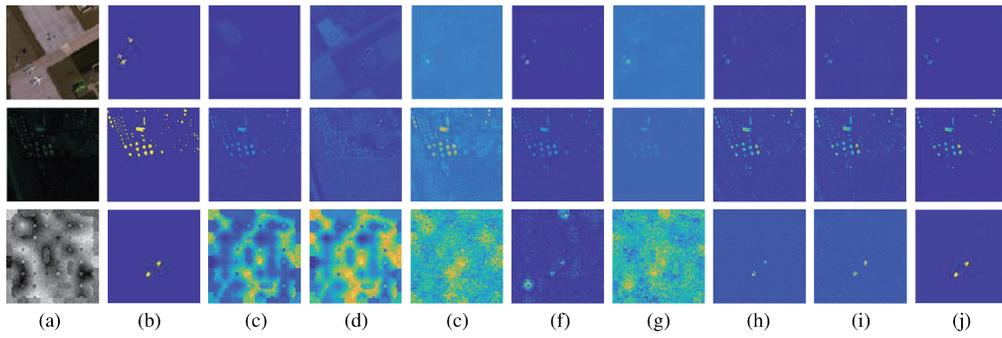


图3 三种数据集检测图不同方法对比

Fig. 3 Comparison of different methods for detecting graphs on three data sets

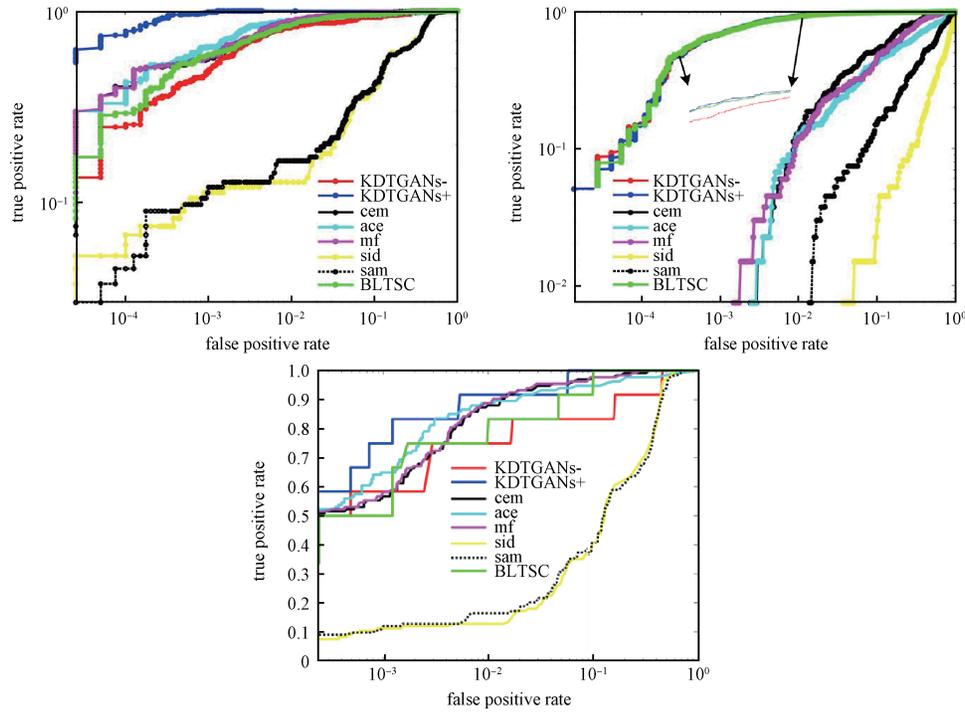


图4 Sandiego、EI Segundo和Synthetic数据集在不同算法上的ROC曲线

Fig. 4 ROC curves of different algorithms on Sandiego, EI Segundo and Synthetic

表1 不同检测算法在Sandiego、EI Segundo和Synthetic数据集的AUC值

Table 1 AUC values of Sandiego, EI Segundo and Synthetic data sets in different detection algorithms

HSI	AUC values(P_D , P_F)		
	Sandiego	EI Segundo	Synthetic
CEM	0.990 4	0.829 8	0.979 5
ACE	0.978 8	0.837 7	0.969 9
MF	0.990 7	0.748 7	0.975 4
SID	0.811 9	0.630 6	0.552 1
SAM	0.806 1	0.591 8	0.623 8
BLTSC	0.985 1	0.983 3	0.993 1
KDTGANs-	0.975 2	0.981 9	0.986 5
KDTGANs+	0.999 9	0.992 6	0.999 4
KDTGAN	0.999 9	0.993 0	0.999 6

公式(8)中, P_{signal} 表示数据的功率, P_{noise} 表示噪声的功率。

3.2.2 消融实验

为评估提出的模型各模块的有效性,对教师模型KDTGAN和学生模型KDTGAN+进行了消融实验,采用弗雷谢起始距离(Fréchet Inception Distance, FID)分数来衡量GAN训练过程的重构精度受本文提出的模块的影响程度,并对不同模块减去后的检测精度进行对比,结果如表2所示。针对学生模型,如3.2.1中展示的KDTGANs+和KDTGANs-以证明本文提出的蒸馏损失,大大提高了学生模型的检测精度。

表2 KDTGAN模型在Sandiego数据集消融实验
Table 2 KDTGAN model ablation experiment on Sandiego dataset

模型	FID	AUC
KDTGAN- L_{STT}	38.61	0.991 7
KDTGAN-多尺度	40.98	0.986 5
KDTGAN	39.19	0.999 9

4 结束语

本研究提出了一种基于知识蒸馏的轻量级 Transformer-GAN 模型, 重点关注高光谱目标检测。通过设计生成器结构, 旨在使模型更专注于捕获高光谱图像中的全局信息, 并构建一个损失函数来应对背景训练样本不纯的挑战。通过引入蒸馏损失, 巧妙地应用了知识蒸馏策略, 从光谱和空间角度促进了高层次信息向学生模型的转移, 达到计算量降低的同时, 重构纯背景能力不减, 并实现高光谱目标检测。实验结果表明: 在高光谱目标检测方面, 本文提出的 KDTGAN 模型明显优于其他检测器。

参考文献

- [1] Microimages. Introduction to hyperspectral imaging[EB/OL]. [2024-01-19] <https://www.microimages.com/documentation/Tutorials/hyprspec.pdf>.
- [2] FARRAND W H, HARSANYI J C. Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique[J]. Remote Sensing of Environment, 1997, 59(1): 64-76.
- [3] LI W, DU Q, ZHANG B. Combined sparse and collaborative representation for hyperspectral target detection[J]. Pattern Recognition, 2015, 48(12): 3904-3916.
- [4] KRAUT S, SCHARF L L, MCWHORTER L T. Adaptive subspace detectors[J]. IEEE Transactions on Signal Processing, 2001, 49(1): 1-16.
- [5] FUHRMANN D R, KELLY E J, NITZBERG R. A CFAR adaptive matched filter detector[J]. IEEE Transactions on Aerospace and Electronic Systems, 1992, 28(1): 208-216.
- [6] 赵佳乐, 王广龙, 周冰, 等. 深度学习下高光谱图像目标检测技术研究进展[J]. 激光杂志, 2022, 43(10): 1-6. DOI:10.14016/j.cnki.jgzz.2022.10.001.
- [7] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems, December 8-13, 2014, Montreal Canada. Cambridge: MIT Press, 2014.
- [8] KINGMA D P, WELLING M. Auto-encoding variational Bayes[EB/OL]. arXiv preprint, 2013 arXiv: 1312.6114.
- [9] XIE Weiyang, ZHANG Xin, LI Yunsong, et al. Background learning based on target suppression constraint for hyperspectral target detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 5887-5897.
- [10] SHEN Kun, XIE Weixin, TANG Haojin, et al. Dense convolution siamese network for hyperspectral image target detection[C]//2022 5th International Conference on Information Communication and Signal Processing (ICICSP), November 26-28, 2022, Shenzhen, China. IEEE, 2022.
- [11] CHEN Xi, WANG Yulei, CHE Zongwei, et al. Contrastive learning for hyperspectral target detection[C]//IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, July 17-22, 2022, Kuala Lumpur, Malaysia. IEEE, 2022.
- [12] XIE Weiyang, YANG Jian, LEI Jie, et al. SRUN: Spectral regularized unsupervised networks for hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 58(2): 1463-1474.
- [13] 白凯玮. 基于分布式深度学习的高光谱图像目标检测算法研究[D]. 西安: 西安电子科技大学, 2022. DOI: 10.27389/d.cnki.gxadu.2022.002466.
- [14] QIN Haonan, XIE Weiyang, LI Yunsong, et al. PTGAN: A proposal-weighted two-stage GAN with attention for hyperspectral target detection[C]//2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, July 11-16, 2021, Brussels, Belgium. IEEE, 2021.
- [15] 鲍亮. 基于深度学习的高光谱遥感图像目标检测算法[D]. 南京: 南京理工大学, 2019. DOI:10.27241/d.cnki.gnjgu.2019.002211.
- [16] 解静. 基于深度学习的高光谱图像目标检测算法研究[D]. 西安: 西安电子科技大学, 2019. DOI: 10.27389/d.cnki.gxadu.2019.000351.
- [17] KIM J H, KIM J, JOUNG J. Siamese hyperspectral target detection using synthetic training data[J]. Electronics Letters, 2020, 56(21): 1116-1118.
- [18] QIN Haonan, XIE Weiyang, LI Yunsong, et al. HTD-VIT: Spectral-spatial joint hyperspectral target detection with vision transformer[C]//IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Sympos-

- sium, July 17-22, 2022, Kuala Lumpur, Malaysia. IEEE, 2022.
- [19] JIAO Jinyue, GONG Zhiqiang, ZHONG Ping. Triplet spectral-wise transformer network for hyperspectral target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5519817.
- [20] SHEN Dunbin, MA Xiaorui, KONG Wenfeng, et al. Hyperspectral target detection based on interpretable representation network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5519416.
- [21] 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.
- [22] LI Yunsong, QIN Haonan, XIE Weiyang. HTDFormer: Hyperspectral target detection based on transformer with distributed learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5524715.
- [23] KARRAS Tero, LAINE Samuli, AILA Timo. A style-based generator architecture for generative adversarial networks[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. IEEE, 2019.
- [24] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[EB/OL]. <https://doi.org/10.48550/arXiv.1706.03762>.
- [25] PARK W, KIM D, LU Y, et al. Relational knowledge distillation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15-20, 2019, Long Beach, CA, USA. IEEE, 2019: 3967-3976.

[作者简介]

谢 雯 1989年生, 副教授, 研究生导师。

闪晨超 1999年生, 硕士研究生。

张哲哲 1997年生, 硕士研究生。

张嘉鹏 1998年生, 硕士研究生。

(本文编辑: 傅 杰)