

显式鉴别驱动-未知类自动聚类的开放世界半监督学习方法

宋佳伦^① 杜兰^{*①} 陈健^①

^①(西安电子科技大学雷达信号处理全国重点实验室 西安 710071)

摘要: 传统目标识别常基于闭集假设, 即测试目标类别均属训练集。然而真实世界具有开放性, 除训练集已知类外, 识别任务还关注未知新目标的解译, 要求模型兼具已知类识别与未知类发现并聚类的能力。针对上述问题, 该文提出一种显式鉴别驱动-未知类自动聚类的直推式开放世界半监督学习方法, 综合利用少量标记已知类训练样本与大量无标记待测已知类/未知类样本学习模型, 实现已知类识别与未知类聚类。所提方法重点包含基于已知类边界极值分布动态扩展的无标记已知类-未知类鉴别模块, 结合极值理论建模标记已知类边界分布, 并在半监督学习中对该分布进行动态完善, 提升模型鉴别能力; 对所鉴别高置信未知类样本, 基于近邻交并比关系合并的未知类自动聚类模块进行近邻聚类簇合并, 实现未知类自动聚类。鉴别模块和未知类自动聚类模块迭代优化。基于光学CIFAR-10与实测雷达数据的实验表明所提方法具有良好的已知类识别与未知类聚类性能。

关键词: 目标识别; 开放世界半监督学习; 未知类鉴别; 自动聚类; 极值理论

中图分类号: TP181

文献标识码: A

文章编号: 1009-5896(2026)00-0001-13

DOI: 10.11999/JEIT251291

CSTR: 32379.14.JEIT251291

1 引言

目标识别是基于传感器获取的数据识别目标的技术, 能够实现目标类别的判定, 在军用和民用领域均起着十分重要的作用。近年来, 基于深度学习等新兴技术的目标识别研究取得革命性突破, 在理论研究中的性能已几乎达到甚至超越了人类水平^[1]。然而, 现有目标识别研究大都局限于闭集静态假设, 即认为测试阶段所有待识目标类均包含在训练阶段已知类别中。与之相反, 真实世界具有开放性, 已知训练类别分布外的未知新目标类广泛存在。闭集目标识别方法会将未知新目标类强制错判为已知类, 造成严重错误。

针对真实世界开放条件下的目标识别问题, 开集识别(Open-Set Recognition, OSR)首先被提出, 旨在训练阶段基于标记已知类数据学习开集识别模型, 在测试阶段实现已知类的识别与未知类的拒判, 已在安防、监控等多个领域得到广泛应用。近年来, 开放世界半监督学习(Open-World Semi-Supervised Learning, OWSSL)^[2]逐渐引起关注, 旨在训练阶段基于标记已知类数据, 无标记已知类、未知类数据学习模型, 在测试阶段实现已知类的识别与未知类的聚类, 对识别任务关注的未知新目标解译具有重要支撑意义。现有OWSSL研究分为归纳

式与直推式两种。其中, 归纳式OWSSL^[3,4]从标记、无标记训练数据中学得模型并应用于测试环境的已知类识别/未知类聚类。随着开放条件目标识别的应用场景日益丰富、目标类型日益繁杂, 仅依赖训练阶段预先采集的局限类别数据所学模型常与测试环境类别严重失配, 已越来越难以满足识别任务需要。相应地, 直推式OWSSL^[2]被提出, 旨在利用标记已知类训练数据和大量无标记已知类/未知类测试数据学习模型, 提升模型已知类识别能力的同时直接得到测试环境未知类的聚类结果(如图1所示)。本文亦采取直推式OWSSL范式, 并与所有OWSSL方法在直推式学习模式下进行对比。

目前, OWSSL研究已取得一定进展。在未知类聚类方面, 区别于传统聚类算法聚类标准不一(例, 颜色、形状)、难以保证聚类结果和实际语义类别相符, OWSSL利用已知类数据学习语义类别聚合知识, 帮助模型实现无标记未知类数据的准确聚类。此外, 归纳式/直推式OWSSL仅存在参与模型学习数据上的差异, 二者的方法分类和解决的核心问题一致, 以下按照类别聚合知识学习方式的不同, 对现有OWSSL方法进行统一介绍:

基于相似度学习的方法: 通过同类正样本对高相似度学习获取类别聚合知识。文献^[2]约束标记已知类同类样本、无标记数据近邻样本具有高相似度表示, 引导无标记已知类向标记已知类分布聚合, 无标记未知类近邻样本不断聚合至预设聚类簇个数, 并约束无标记样本聚类概率均衡以避免其聚类为单一簇的模式崩塌。考虑到标记已知类监督学习相比无标记数据聚类学习收敛较快, 导致未知类特

收稿日期: 2026-xx-xx; 改回日期: 2026-06-24; 网络出版: 2026-07-04

*通信作者: 杜兰 dulan@mail.xidian.edu.cn

基金项目: 国家自然科学基金(U24B20137, U21B2039, 62201433)

Foundation Items: The National Natural Science Foundation of China (U24B20137, U21B2039, 62201433)

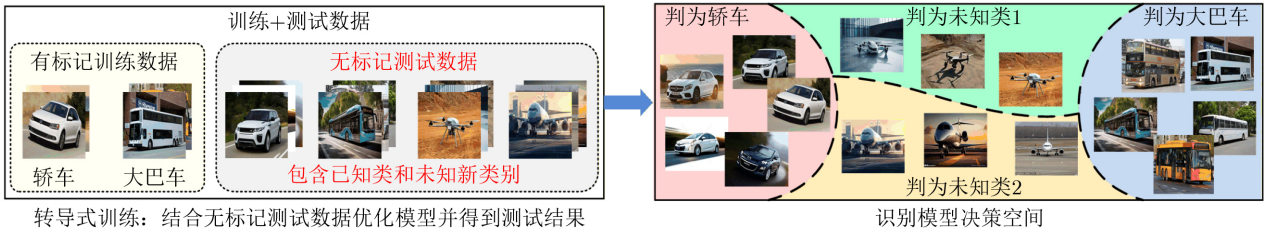


图1 直推式OWSSL示意图

征表示易偏向已知类,文献[3]挑选高置信无标记未知类样本加入监督训练以平衡二者学习进程。

基于隐空间学习的方法:通常将已知类识别和未知类聚类表示为一个 $K+U$ 类联合分类任务(其中 K 为已知类类别数, U 为未知类预设类别数),标记/无标记样本在联合分类任务的特征/分类空间中共享类别聚合知识,有利于缓解相似度方法中收敛进程难平衡的问题[4]。文献[4]对标记/无标记数据统一施加 $K+U$ 交叉熵损失,其中无标记数据结合Sinkhorn-Knopp算法[5]生成 $K+U$ 类分配均匀的伪标签,避免模式崩塌。为提高已知类聚合知识对未知类聚类的可迁移性,文献[6]引入标记已知类数据-无标记数据混合策略。文献[7]为提升模型已知类识别能力,挑选无标记样本中的疑似已知类并加强其向已知类分布的学习;文献[8]设计了一种即插即用的目标子区域注意力聚焦策略,可整合入现有OWSSL方法提升各类间可分性。

综上,现有OWSSL方法主要基于相似度学习或隐空间学习的途径,从标记已知类数据中学习类别聚合知识,引导无标记已知类向标记已知类分布聚合,无标记未知类同类样本分布聚合,实现无标记已知类识别与未知类聚类。然而,现有方法尚存在以下不足:**(1)缺乏显式的无标记已知类-未知类鉴别手段:**现有方法利用标记已知类数据学习的类别聚合知识引导所有无标记数据聚合,不对无标记已知类/未知类数据预先区分,致使无标记已知-未知类之间易错判。在未知类表征偏向问题下,无标记未知类错判为已知类的问题尤为显著;**(2)局限于无标记未知类类别个数已知的强假设:**现有OWSSL方法大都假设未知类类别个数已知,并在预设未知类聚类个数的条件下实现未知类聚类,然而未知类类别个数先验在实际应用中往往难以预知,致使预设未知类聚类个数的方法易错误聚类。

针对上述问题,本文提出显式鉴别驱动-未知类自动聚类的开放世界半监督学习方法。在隐空间学习OWSSL框架基础上,引入基于已知类极值分布动态扩展的无标记已知-未知类鉴别模块,结合极值理论(Extreme Value Theory, EVT)拟合标记

已知类极值分布作为鉴别依据,并在半监督学习中动态调整极值分布以提升模型鉴别能力,使**已知类/未知类特征朝不同方向优化以减少错判**;在此基础上,对所鉴别高置信未知类,通过近邻交并比关系合并的自动聚类模块,实现**无标记未知类类别个数的自动估计与样本归属划分**。鉴别模块与聚类模块联合优化:鉴别模块选取高置信未知类以保障未知类聚类模块的准确学习;聚类模块确立未知类聚类簇后与已知类特征分布进行推远,促进已知类-未知类特征鉴别性提升。本文基于公共Benchmark数据CIFAR-10[9]开展理论性能验证;考虑到合成孔径雷达(Synthetic Aperture Radar, SAR)在目标识别中的广泛应用及其本身面向开放条件的典型“非合作”特点,本文还基于实测SAR数据:MSTAR[10]、ATRNet-STAR[11]开展性能验证。实验结果表明所提方法具有先进的已知类识别性能与未知类聚类性能。

2 方法介绍

2.1 方法整体框架与训练总损失

图2展示了所提显式鉴别驱动-未知类自动聚类的OWSSL方法,采用直推式学习模式,基于标记已知类训练数据和无标记已知类/未知类待测数据学习模型,直接得到待测已知类识别结果与未知类聚类结果。考虑到原型网络(Prototypical Network, PN)[12]具有良好的类内紧凑与类间可分特征表示能力,所提方法采用基于原型网络的识别模型(由特征提取网络与类原型构成)。给定标记已知类训练数据 $\mathbf{I}_1 = \{(\mathbf{x}_1^i, y_1^i)\}_{i=1}^q$,其中 q 表示标记已知类训练样本总数, $y_1^i \in \mathcal{C}_1 = \{c_1, c_2, \dots, c_K\}$ 表示样本 \mathbf{x}_1^i 的类别标签, c_K 表示类别 K , K 为已知类类别总数;给定无标记已知类/未知类待测数据 $\mathbf{I}_u = \{\mathbf{x}_u^j\}_{j=1}^m$,其中 m 表示无标记已知类/未知类待测样本总数, $\forall \mathbf{x}_u^j$ 的类别属于集合 $\mathcal{C}_u = \{c_1, \dots, c_K, c_{K+1}, \dots, c_{K+U}\}$, $\mathcal{C}_1 \subset \mathcal{C}_u$ 且 $\mathcal{C}_u - \mathcal{C}_1 = \{c_{K+1}, \dots, c_{K+U}\}$, U 表示未知类类别总数。在基于标记已知类数据完成预热训练后,所提方法训练流程主要分为三步:

(1)无标记已知-未知类鉴别:输入标记已知类

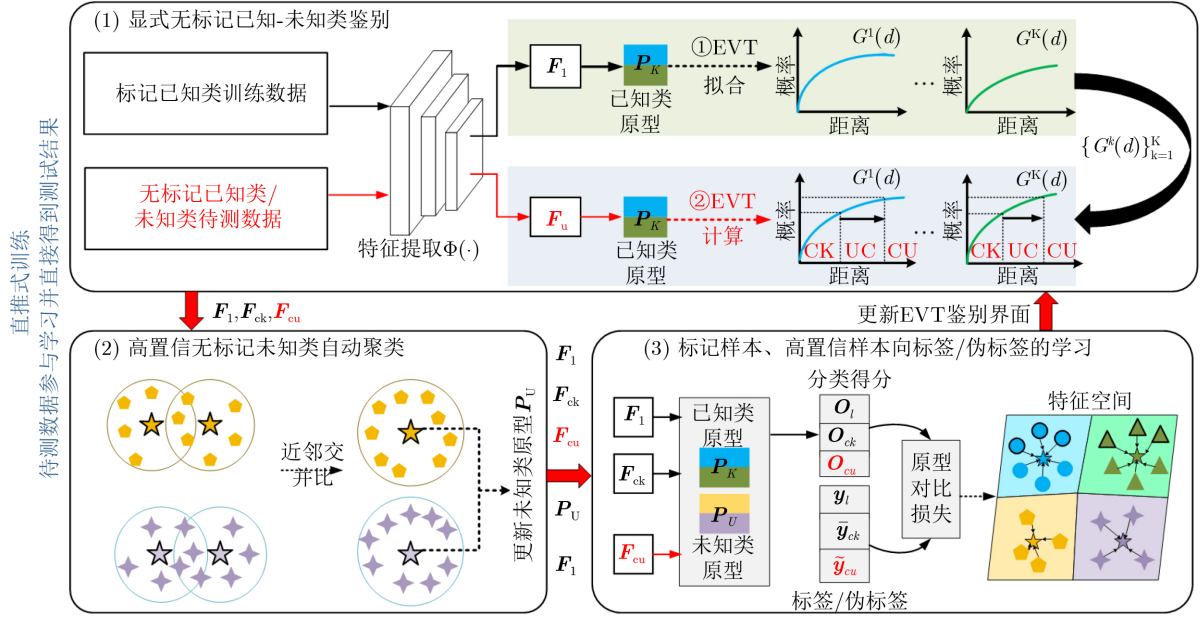


图 2 显式鉴别驱动-未知类自动聚类的OWSSL方法示意图

训练数据 I_1 与无标记已知类/未知类待测数据 I_u ，经特征提取网络 $\Phi(\cdot)$ 得到对应特征 F_1 与 F_u ，计算 F_1 各类特征均值以确立已知类原型 $P_K = \{p_k\}_{k=1}^K$ ，随后执行①EVT拟合：对 F_1 各类特征与其对应类别原型的距离进行极值拟合以确立各已知类分布边界 $\{G^k(x)\}_{k=1}^K$ [13]，②EVT计算：将 F_u 与各已知类原型距离代入 $\{G^k(x)\}_{k=1}^K$ 实现高置信已知类(Confident Knowns, CK) F_{ck} 与高置信未知类(Confident Unknowns, CU) F_{cu} 的鉴别；

(2)高置信无标记未知类自动聚类：输入 F_{cu} ，首先以未知类个数初值 U_0 的K-means对 F_{cu} 进行聚类初始化，并基于各聚类簇近邻交并比合并冗余聚类簇、更新未知类原型 $P_U = \{p_u\}_{u=K+1}^{K+U}$ 。特别地，由于未知类个数真值 U^* 未知，实验设定 U_0 为一较大值，并在聚类模块优化下实现向实际未知类个数 U^* 的不断逼近；

(3)标记样本、高置信样本向标签/伪标签的学习：输入 F_1, F_{ck}, F_{cu} ，已知类原型 P_K 及步骤(2)中更新的 P_U ，结合标签/伪标签在原型对比损失约束下端到端优化特征提取网络 $\Phi(\cdot)$ 与 $[P_K, P_U]$ ，实现 F_1, F_{ck}, F_{cu} 各类特征与本类原型相近、与异类原型相远。

上述步骤(1)~(3)在模型训练中不断迭代，以显式鉴别模块为驱动，为聚类模块挑选高置信未知类进行聚类以保障未知类聚类精度，随后与标记、高置信已知类通过标签/伪标签的学习相互推远以提升二者间可分性，将高置信已知类加入鉴别模块 EVT拟合以进一步完善鉴别边界，形成学习闭环。最终进行待测样本特征与已知类原型和未知类

原型 $[P_K, P_U]$ 间的距离度量，按最近距离原则得到待测样本已知类识别/未知类聚类结果。结合原型对比学习的训练总损失如下：

$$\begin{aligned} L_{\text{total}} &= L_{\text{PL}}(\mathbf{x}_1) + L_{\text{PL}}(\mathbf{x}_{ck}) + L_{\text{PL}}(\mathbf{x}_{cu}) \\ &= -\mathbf{y}_1 \cdot \log(\text{softmax}(\mathbf{o}_1)) \\ &\quad -\bar{\mathbf{y}}_{ck} \cdot \log(\text{softmax}(\mathbf{o}_{ck})) \\ &\quad -\tilde{\mathbf{y}}_{cu} \cdot \log(\text{softmax}(\mathbf{o}_{cu})) \end{aligned} \quad (1)$$

其中 $\mathbf{y}_1 = [y_1^1, \dots, y_1^K, 0, \dots, 0] \in \mathbb{R}^{K+U}$ ， $\bar{\mathbf{y}}_{ck} = [y_{ck}^1, \dots, y_{ck}^K, 0, \dots, 0] \in \mathbb{R}^{K+U}$ 分别对应标记样本 \mathbf{x}_1 ，高置信已知类样本 \mathbf{x}_{ck} 的one-hot类别标签，在标记/高置信已知类别维度置1，其余维度置0，引导标记/高置信已知类样本向对应已知类分布聚合； $\tilde{\mathbf{y}}_{cu} = [0, \dots, 0, \tilde{y}_{cu}^1, \dots, \tilde{y}_{cu}^U] \in \mathbb{R}^{K+U}$ 为高置信未知类样本 \mathbf{x}_{cu} 的伪标签，前 K 维整体置0，后 U 维对应 U 类未知类聚类概率伪标签。每个样本分类得分 $\mathbf{o} = d(\mathbf{f}, [P_K, P_U]) = d(\mathbf{f}, [p_1, \dots, p_K, p_{K+1}, \dots, p_{K+U}])$ 。此外，与现有基于隐空间学习的方法一致，未知类聚类概率伪标签基于Sinkhorn-Knopp [5] 算法求解，表达式如下：

$$\begin{aligned} \min_{\mathbf{Y} \in \mathbb{R}^{B \times U}} \sum_b \left(\sum_{i=1}^U \tilde{y}_{cu,b}^i - 1 \right) + \sum_i \left(\sum_{b=1}^B \tilde{y}_{cu,b}^i - \frac{B}{U} \right) \\ - \text{tr}(\tilde{\mathbf{Y}}^T \mathbf{O}), \\ \mathbf{O} = [\text{softmax}(\mathbf{o}_{cu,1}), \dots, \text{softmax}(\mathbf{o}_{cu,B})] \in \mathbb{R}^{B \times U} \end{aligned} \quad (2)$$

给定 B 个高置信未知类样本， $\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}_{cu,1}, \dots, \tilde{\mathbf{y}}_{cu,b}, \dots, \tilde{\mathbf{y}}_{cu,B}] \in \mathbb{R}^{B \times U}$ 表示其聚类概率伪标签， $\forall \tilde{\mathbf{y}}_{cu,b} = [\tilde{y}_{cu,b}^1, \dots, \tilde{y}_{cu,b}^i, \dots, \tilde{y}_{cu,b}^U] \in \mathbb{R}^U$ ， $\text{tr}(\cdot)$ 表示矩阵

迹运算。式(2)第一项保证每个样本聚类伪标签概率和等于1；第二项保证B个样本在U类任意类别上的伪标签元素和均等于B/U，避免聚类模式崩塌；第三项最大化矩阵 $\hat{Y}^T O$ 的迹以生成比当前预测聚类概率分布更尖锐的伪标签，引导样本特征向对应聚类簇聚合。

步骤(1)、(2)对应两模块是本文方法的核心，下面对步骤(1)、(2)两模块进行详细介绍。

2.2 基于已知类极值分布动态扩展的显式无标记已知类-未知类鉴别模块

给定标记已知类样本特征集 $F_1 = \{f_1^i\}_{i=1}^q$ ，无标记样本特征集 $F_u = \{f_u^j\}_{j=1}^m$ ，该鉴别模块详细介绍如下：

①EVT拟合：建立已知类-未知类鉴别能力。在识别模型每轮训练优化前，基于 F_1 与高置信已知类样本特征集 F_{ck} 拟合EVT(基于标记已知类样本 I_1 预热训练后，首次拟合EVT仅依赖 F_1 ，待挑选出高置信已知类后再将 F_{ck} 加入EVT拟合)，首先选取各标记/高置信已知类特征与对应类别原型的距离极大值^[14]：

$$\begin{aligned} D_{kno}^k &= \text{sort} \{d_1^k, \dots, d_n^k\}, \forall d_i^k = d(f_k^i, p_k), \\ k &\in \{1, \dots, K\}, i \in \{1, \dots, n\}, \\ (D_{kno}^k)^{EVT} &= \{d_{n-\Delta}^k, \dots, d_n^k\} \end{aligned} \quad (3)$$

其中， D_{kno}^k 表示第 k 类标记/高置信已知类特征 $\{f_k^i\}_{i=1}^n \in \{F_1, F_{ck}\}$ 与第 k 类已知类原型 p_k 从大到小排序的距离集合(注：本文使用余弦相似度 $d(\cdot)$ 度量，故选取排序后距离序列尾部余弦相似度极小值进行极值拟合)， n 表示 $\{F_1, F_{ck}\}$ 中第 k 类已知类特征 $\{f_k^i\}_{i=1}^n$ 数量， sort 表示排序函数， Δ 为各类用于拟合EVT的极值数量(实验选取常用极值比例 $\Delta = 20\% \times n$)， $(D_{kno}^k)^{EVT}$ 表示选取用于第 k 类已知类用于EVT拟合的极值距离集合(如图3(a)所示)。随后将 $(D_{kno}^k)^{EVT}$ 用于拟合EVT累积分布函数：

$$G(d) = 1 - \left(1 + \frac{\xi \times d}{\beta}\right)^{-1/\xi} \quad (4)$$

其中 $G(d)$ 为广义帕雷托EVT累积分布函数， ξ 和 β 分别表示分布函数形状参数和尺度参数。对所有 K 类已知类执行上述操作得到 $\{G^1(d), \dots, G^K(d)\}$ ，完成EVT拟合。每一类EVT累积分布函数 $G^k(d)$ 的横坐标对应随机变量： $d(f, p_k)$ (任一特征 f 与第 k 类已知类原型 p_k 的距离)，纵坐标对应累积概率(如图3(b)所示)。 $d(f, p_k) \in [0, +\infty)$ ，累积概率取值 $\in [0, 1]$ 。累积概率越大，表明 f 离当前拟合边界越近，反之越远。

②EVT计算：旨在结合 $\{G^1(d), \dots, G^K(d)\}$ 实现 F_u 中高置信已知类/高置信未知类/不确定类(Unconfident classes, UC)的判定。计算 F_u 特征与各已知类原型的距离：

$$\begin{aligned} D_{umo}^k &= \{\hat{d}_1^k, \dots, \hat{d}_m^k\}, \forall \hat{d}_j^k = d(f_u^j, p_k), \\ k &\in \{1, \dots, K\}, f_u^j \in F_u \end{aligned} \quad (5)$$

其中 D_{umo}^k 表示所有无标记样本特征 $F_u = \{f_u^j\}_{j=1}^m$ 与第 k 类已知类原型 p_k 的距离集合。将 D_{umo}^k 代入 $G^k(d), k \in \{1, \dots, K\}$ ，即可基于下式实现 f_u^j 对应CK/UC/CU的判断：

$$\begin{aligned} \hat{c}_u^j &\in \text{CK}_k \text{ if } G^k(\hat{d}_j^k) \leq \tau_1, \\ \hat{c}_u^j &\in \text{CU} \text{ if } \forall k, G^k(\hat{d}_j^k) \geq \tau_2, \\ \hat{c}_u^j &\in \text{UC} \text{ if } \tau_1 < G^k(\hat{d}_j^k) < \tau_2 \end{aligned} \quad (6)$$

其中， \hat{c}_u^j 对应第 j 个无标记样本的CK/UC/CU判断结果。考虑到模型训练初期仅少量标记已知类拟合边界难以表征完整已知类边界，且无标记已知类-未知类可分性不足，未知类易侵入已知类边界，实验中训练初期选取较严苛的高置信已知类判定阈值 $\tau_1 = 0.5$ 、未知类判定阈值 $\tau_2 = 0.9$ 。

● 当 $G^k(\hat{d}_j^k) \leq 0.5$ ，表明 \hat{d}_j^k (f_u^j 与 p_k 的距离)在

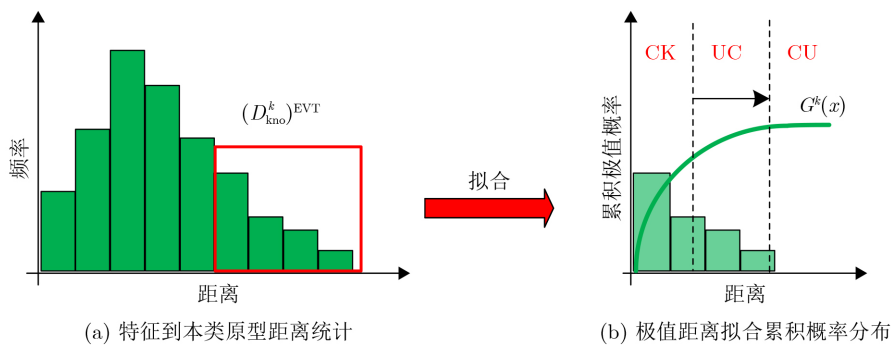


图3 EVT拟合示意图

多数第 k 类标记/高置信已知类特征到 p_k 距离极值的范围内，即 f_{ii}^j 对应无标记样本 x_{ii}^j 有高置信度属于已知类 k (以 CK_k 表示)；

- 当 $\forall k, G^k(\hat{d}_j^k) \geq 0.9$ ，任一已知类 EVT 分布中 \hat{d}_j^k 附近累积概率变化已趋于平缓并接近 1，表明 f_{ii}^j 与任一已知类原型 p_k 距离，远超过对应已知类标记/高置信样本特征到其原型距离的极值，即与各已知类分布距离均较远， x_{ii}^j 有高置信度属于未知类；

- 当处于二者中间情况 $0.5 < G^k(\hat{d}_j^k) < 0.9$ ，则暂不确定其类别。

随模型训练中更多无标记已知类样本加入 EVT 拟合，各已知类分布趋于完备，且无标记已知类-未知类可分性逐渐提升。相应地，式(6)中 τ_1 随训练轮次由严苛的初值 0.5 逐渐上升至 0.9，最终保证所有无标记样本均得到属于已知类/未知类的判断，并在图 2 步骤(3)学习中朝不同方向优化，减少已知-未知类间混淆。

2.3 基于近邻交并比关系合并的未知类自动聚类模块

图 4 给出了所提无标记未知类自动聚类模块示意图。给定前述模块已鉴别的高置信无标记未知类特征集合 $F_{cu} = \{F_{cu}^{K+1}, \dots, F_{cu}^{K+U}\}$ ，包含 U 个聚类簇 (以初值 $U = U_0$ 的 K-means 聚类进行初始化，并基于自动聚类模块在模型迭代优化中不断向真值聚类个数 U^* 缩减，每次迭代自动聚类完成后以新聚类中心作为未知类原型 P_U 对其进行反馈更新)。对 $\forall F_{cu}^{K+i}, i \in \{1, \dots, U\}$ ，统计其所含特征 $f_{cu,j}^{K+i} \in F_{cu}^{K+i}$ 与所有未知类聚类簇原型 $P_U = [p_{K+1}, \dots, p_{K+U}]$ 的距离并进行排序：

$$\begin{aligned} & d(f_{cu,j}^{K+i}, P_U) \\ &= \left[(d_{cu,j}^{K+i})^{K+1}, \dots, (d_{cu,j}^{K+i})^{K+u}, \dots, (d_{cu,j}^{K+i})^{K+U} \right], \\ & \text{sort}(d(f_{cu,j}^{K+i}, P_U)) \\ &= \left[(d_{cu,j}^{K+i})^{K+i}, (d_{cu,j}^{K+i})^{K+j}, \dots, (d_{cu,j}^{K+i})^{K+U} \right] \quad (7) \end{aligned}$$

其中 $(d_{cu,j}^{K+i})^{K+u}$ 表示第 $K+i$ 类聚类簇特征 $f_{cu,j}^{K+i}$ 与第 $K+u$ 类未知类聚类簇原型 p_{K+u} 之间的余弦相似度距离，红色标出部分表示 $f_{cu,j}^{K+i}$ 与本类原型 p_{K+i} 之间相似度最高，且与非本类的近邻原型 p_{K+j} 之间相似度次高。对 $F_{cu} = \{F_{cu}^{K+1}, \dots, F_{cu}^{K+U}\}$ 所有特征执行式(7)。若 F_{cu}^{K+i} 所含特征除与本类 p_{K+i} 距离最近外，大部分特征的近邻异类原型均为 p_{K+j} ，且 F_{cu}^{K+j} 大部分特征的近邻异类原型同时也为 p_{K+i} ，则二者存在较大重叠应合并：

$$\text{IOU}_{K+i,K+j} = \frac{|(f_{cu}^{K+i})^{K+j} \cap (f_{cu}^{K+j})^{K+i}|}{|(f_{cu}^{K+i})^{K+j} \cup (f_{cu}^{K+j})^{K+i}|},$$

$$i, j \in \{1, \dots, U\}$$

if $\text{IOU}_{K+i,K+j} > \tau_c$ ，合并 f_{cu}^{K+i} 与 f_{cu}^{K+j} (8)

其中 $(f_{cu}^{K+i})^{K+j}$ 表示聚类簇 F_{cu}^{K+j} 中近邻异类原型为 p_{K+i} 的特征集合， $\text{IOU}_{K+i,K+j}$ 表示聚类簇 F_{cu}^{K+i} 与 F_{cu}^{K+j} 之间的交并比， τ_c 表示聚合阈值。当 $\text{IOU}_{K+i,K+j} > \tau_c$ 时，将聚类簇 F_{cu}^{K+i} 与 F_{cu}^{K+j} 进行合并，且新聚类簇原型初始化为 F_{cu}^{K+i} 与 F_{cu}^{K+j} 的特征均值。聚合阈值 $\tau_c \in [0.5, 1]$ 并基于下式自动搜索确立：

$$\sum_{f \in \{f_i, f_{ck}, f_{cu}\}} \frac{b(f) - a(f)}{\max\{a(f), b(f)\}} \quad (9)$$

式(9)为 Silhouette cluster validity index 度量用于衡量聚类可分性，其中 $a(f)$ 为特征 f 与其所在类别所有特征距离的均值， $b(f)$ 为特征 f 与其最近邻异类所有特征距离的均值。为保证未知类自动聚类后 $K+U$ 类分类整体最优，且基于标记已知类、高置信已知类学得可分性提升式(9)计算稳定性，本文在上式计算中同时考虑标记已知类特征 f_i ，高置信已知类、未知类特征 f_{ck} 、 f_{cu} 。

随模型迭代优化，前述鉴别模块挑选出更多高

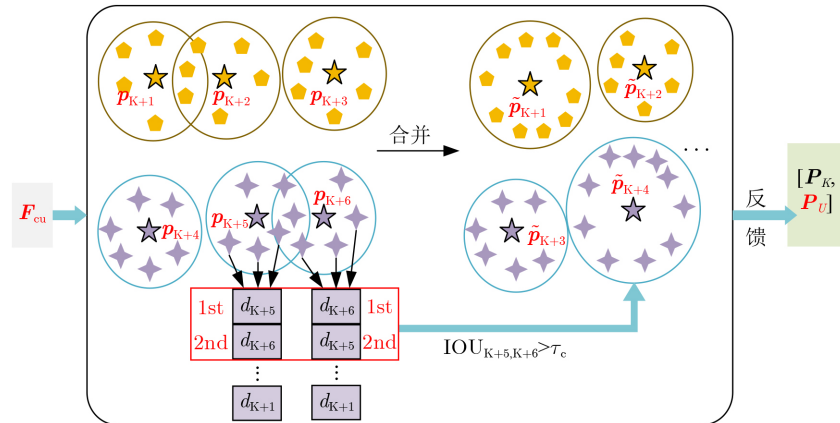


图 4 基于近邻交并比关系合并的未知类自动聚类模块示意图

置信未知类样本参与学习, 自动聚类模块将基于未知类特征-聚类簇原型近邻关系实现初始 U_0 类到逼近真值 U^* 类聚类的优化, 最终实现对无标记样本中所有未知类的自动聚类, 摆脱现有方法对“未知类个数已知”强假设的依赖。

3 实验与分析

3.1 实验数据与设置

3.1.1 实验数据

本文结合OWSSL领域Benchmark数据集CIFAR-10^[9], 以及实测SAR数据集: MSTAR^[10]、ATRNet-STAR^[11]开展实验验证。其中, CIFAR-10包含飞机、小汽车、鸟、猫、鹿、狗、蛙、马、羊、卡车共10类光学RGB图像, 每张图像尺寸 32×32 , 共包含60 000张图像, 其中50 000张训练图像(每类5 000个训练样本), 10 000张测试图像(每类1 000个测试样本); MSTAR包含BMP2、BTR70、T72、BTR60、2S1、BRDM2、D7、T62、ZIL131和ZSU234共10类军事车辆目标SAR图像, 每张图像尺寸 128×128 , 共包含4 873张图像, 其中2 447张训练图像(每类232~299个训练样本不等), 2 426张测试图像(每类195~274个测试样本不等); ATRNet-STAR是2026年最新发布的大型SAR目标识别数据集, 包含各型轿车、SUV、公共汽车、卡车、救护车、工程车共40型民用车辆目标SAR图像, 每张图像尺寸 128×128 , 标准评估设置下共包含97 575张图像, 其中68 091张训练图像(每类1 225~2 342个训练样本不等), 29 284张测试图像(每类464~1 039个测试样本不等), 可支撑少量标记数据与大量无标记数据实验条件下的OWSSL验证。

3.1.2 实验设置

(1)OWSSL实验数据划分

OWSSL面向仅少量已知类别数据及时标注, 大量无标记已知类/未知类数据待判别的实际开放条件识别环境。为模拟该实验条件, 须对原始实验数据作以下划分:

针对CIFAR-10数据集, 遵从计算机视觉领域OWSSL研究的通用设置: 随机选取5类作为已知类(从原始训练数据随机采样10%作为标记已知类, 原始测试数据作为无标记已知类), 余下5类作为未知类(原始测试数据作为无标记未知类); MSTAR数据集划分方式与上述CIFAR-10划分保持一致; ATRNet-STAR数据集设置: 随机选取30类作为已知类(从原始训练数据随机采样30%作为标记已知类, 余下训练数据和原始测试数据合并作为无标记已知类), 余下10类作为未知类(原始训练与测试数据合并作为无标记未知类)。

(2)模型及超参数设置

针对CIFAR-10数据集实验, 以18层ResNet作为原型识别模型的特征提取网络。已知类原型 P_K 以各标记已知类特征均值进行初始化; 以聚类个数 $U_0 = 20$ 的K-means对所鉴别高置信未知类进行聚类初始化, 并以相应聚类中心作初始未知类原型。训练总轮数200(前10轮仅标记已知类预热训练, 11至200轮加入无标记样本进行OWSSL学习, 其中11至60轮自动聚类学至合理聚类个数 U^* , 61至200轮基于确立的 U^* 学至收敛), 优化器为动量SGD, 初始学习率0.1(80轮, 150轮乘以0.1衰减)。batch size大小256。各已知类用于极值拟合的样本数 Δ 占本类样本的20%, 高置信已知类判别阈值初值为 $\tau_1 = 0.5$, 并按 $0.5 + (0.9 - 0.5) * \sqrt{H/H_{\max}}$ 动态调整, 其中 H, H_{\max} 分别表示当前训练轮次与总训练轮数, 高置信未知类判别阈值设置为 $\tau_2 = 0.9$ 。未知类自动聚类阈值 $\tau_c \in [0.5, 1]$, 并基于式(9)搜索确立。针对MSTAR、ATRNet-STAR数据集实验, batch size大小为64, 且考虑到ATRNet-STAR数据目标类繁杂、已知类-未知类更易混淆, 高置信已知类判别阈值初值为更严苛的 $\tau_1 = 0.2$, 其余设置与CIFAR-10实验设置保持一致。

3.2 评价指标

OWSSL领域通常采用已知类识别准确率 ACC_{known} 和未知类聚类准确率 ACC_{cluster} 两种指标, 定义如下:

$$ACC_{\text{known}} = \frac{1}{N_{\text{known}}} \sum_{i=1}^{N_{\text{known}}} \chi(y_{k,i} = \hat{y}_{k,i}),$$

$$ACC_{\text{cluster}} = \max_{\text{perm} \in \Omega} \frac{1}{N_{\text{unknown}}} \sum_{j=1}^{N_{\text{unknown}}} \chi(y_{u,j} = \text{perm}(\hat{y}_{u,j})) \quad (10)$$

其中 $\chi(\cdot)$ 表示指示函数, 在括号内式成立时取1, 否则取0; $y_{k,i}$ 表示第 i 个无标记已知类样本 $\mathbf{x}_{k,i}$ 的真实类别, $\hat{y}_{k,i}$ 表示其预测类别, N_{known} 为无标记已知类样本总数; $y_{u,j}$ 表示第 j 个无标记未知类样本 $\mathbf{x}_{u,j}$ 的聚类类别, $\hat{y}_{u,j}$ 表示其预测类别, N_{unknown} 为无标记未知类样本总数, $\text{perm}(\cdot) \in \Omega$ 为换序函数, 由于聚类不存在固定的真实类别归属, 现有方法为衡量聚类精度常基于匈牙利算法对 $\{y_{u,j}\}_{j=1}^{N_{\text{unknown}}}$ 和 $\{\hat{y}_{u,j}\}_{j=1}^{N_{\text{unknown}}}$ 进行换序匹配以找到最优对应关系, 保证最优匹配下同类样本被分为同一聚类簇的比例最高, 并在该匹配下计算聚类准确率。

此外, 本文还利用未知类拒判率(Rejection Accuracy on Unknown Samples, RAUS)指标, 定义如下:

$$RAUS = \frac{1}{N_{\text{unknown}}} \sum_{j=1}^{N_{\text{unknown}}} \chi(y_{u,j} > K) \quad (11)$$

当预测类别 $\hat{y}_{u,j} > K$ (即不判为K类已知类)则拒判成功, 该指标衡量模型对无标记已知-未知类的鉴别能力。

3.3 实验与分析

3.3.1 对比实验

(1)光学图像数据集: CIFAR-10对比实验

OWSSL自2022年被提出已取得一定研究进展, 本文选取6种代表性方法与所提方法对比(含2种相似度学习方法, 以及4种隐空间学习方法), 并以开集识别+K-means作为半监督性能下界。此外, 还加入无监督聚类方法验证OWSSL引入少量标记已知类数据的性能增益。所有对比实验均采用直推式学习模式。上述对比方法详细介绍如下:

(a)无监督聚类BootSC^[15]、SNSCC^[16](下界): BootSC在谱聚类基础上引入sinkhorn-knnp算法, 实现样本相似度矩阵、特征嵌入与聚类分配的端到端优化, SNSCC在对比聚类基础上对正样本按距离远近进行差异化加权拉近, 并通过分阶段微调修正同类样本误判为负样本的问题, 提升聚类精度;

(b)ORCA^[2](基于相似度学习的方法): 约束标记已知类同类样本、无标记近邻样本分类得分一致, 并引入软标签平衡已知、未知类学习, 引入最大熵约束避免聚类模式崩塌, 按最大分类概率确定类别归属;

(c)OpenLDN^[3](基于相似度学习的方法): 采用ORCA方法预训练识别模型, 并按分类概率阈值选择高置信已知类、未知类样本赋予固定类别标签, 再采用传统半监督学习方法完成对余下无标记样本的学习;

(d)TRSSL^[4](基于隐空间学习的方法): 将K类

已知类识别与U类未知类聚类表示为统一的K+U分类任务, 并引入sinkhorn-knnp算法, 结合当前无标记样本分类概率生成各类均匀分配的伪标签以引导所有无标记样本向不同类别簇聚合, 按最大分类概率确定类别归属;

(e)PKOSSL^[6](基于隐空间学习的方法): 在TRSSL基础上对标记、无标记样本Mix-up, 提升已知类对未知类聚类学习的辅助效果;

(f)LeGoGCD^[7](基于隐空间学习的方法): 针对以往方法中最大熵/均匀分配伪标签易弱化模型对无标记已知类的学习问题, 提出局部熵正则化策略, 挑选熵值低的疑似已知类并加强其向已知类分布的学习;

(g)AFGCD^[8](基于隐空间学习的方法): 针对以往方法缺乏对目标像素的关注机制, 而错误关注背景易导致类间混淆问题, 基于标记数据学习各类目标区域重要性, 并引导无标记数据的目标区域学习。

以上方法均在未知类个数已知下设计, 本节首先验证OWSSL相比无监督聚类引入少量标记已知类样本带来的性能增益; 然后将所提方法与OWSSL对比方法在未知类个数已知条件下进行对比, 验证所提已知类-未知类鉴别模块的作用; 最终扩展至未知类个数未知条件, 验证所提未知类自动聚类的意义。

如表1所示, OWSSL方法的 ACC_{known} 与 ACC_{cluster} 整体优于无监督聚类(无监督聚类不存在已知类, 本文仅按OWSSL的已知类/未知类划分分别统计无监督聚类方法性能便于对比), 表明少量标记已知类数据所学语义类别聚合知识对已知类识别与未知类聚类均有增益; OWSSL方法中隐空间学习方法缓解了相似度学习方法中标记数据与无标记数据收敛进程难平衡的问题, 整体性能优于相似度学习方法。

表 1 不同方法在CIFAR-10数据集上的RAUS(%), ACC_{known} (%), ACC_{cluster} (%)(每个指标左侧对应未知类个数已知条件下各方法性能, 右侧对应未知类个数未知条件下各方法性能-三次K-means随机初始化下进行学习的平均结果)及K+U估计个数

方法		RAUS(%)		ACC_{known} (%)		ACC_{cluster} (%)		K+U
无监督 聚类	BootSC	-	-	87.70	76.07	88.92	77.42	[8,8,9]
	SNSCC	-	-	88.24	78.10	89.66	78.35	[8,8,9]
	ORCA	91.71	82.45	88.20	80.02	90.40	81.22	[9,11,11]
开放 世界 半监督 学习	OpenLDN	94.54	86.55	93.40	88.35	93.65	84.14	[9,11,11]
	TRSSL	93.20	87.25	95.15	89.07	92.61	84.67	[9,11,11]
	PKOSSL	94.00	87.11	94.90	89.10	91.80	85.00	[9,11,11]
	LeGoGCD	98.80	92.25	64.64	62.62	98.45	89.42	[10,11,11]
	AFGCD	98.04	91.88	87.85	83.72	97.68	88.31	[10,11,11]
本文方法		97.25	95.33	96.12	93.86	96.40	92.74	[10,10,11]

在未知类个数已知时(指标左侧), 所提方法弥补了现有方法缺乏显式无标记已知类-未知类鉴别手段的缺陷, 在已知类识别准确率 ACC_{known} 与未知类聚类准确率 ACC_{cluster} 上的综合性能显著优于对比方法(未知类拒判率RAUS和聚类准确率 ACC_{cluster} 略低于最新对比方法LeGoGCD与AFGCD, 但二者核心策略对标记样本量要求较高, 在10%已知类标注设置下 ACC_{known} 大幅低于所提方法)。在未知类个数未知时(指标右侧), 先将无标记数据输入预训练模型提取特征并结合K-means算法预聚类(无监督聚类/OWSSL方法基于式(9)分别结合无标记/无标记+标记样本特征搜索最优聚类个数 $K+U$), 再在对比方法框架下学习(保持相同主干模型的对比方法预设 $K+U$ 一致)。因K-means搜索 $K+U$ 个数与实际易存在偏差, 对比方法 ACC_{known} 与 ACC_{cluster} 均大幅下降。所提方法摆脱现有方法对“未知类个数已知”的依赖, RAUS、 ACC_{cluster} 高出对比方法LeGoGCD 3.08%、3.32%, ACC_{known} 高出对比方法PKOSSL 4.76%并更大幅优于其他对比方法。

从特征可视化视角分析, 对比方法在未知类个数未知时缺乏自动聚类能力。当K-means搜索聚类

个数与实际不符时: 个数偏小会导致未知类错分至其他簇引发混叠(见图5(a)); 个数偏大则产生冗余簇无法正确归类(见图5(b)); 即使设置个数与实际相符, 对比方法缺乏已知类-未知类鉴别手段因而仍存在显著混叠(见图5(c)); 本文通过动态建模已知类边界分布以鉴别已知、未知类, 引导二者向不同方向优化减少混叠, 并基于所鉴别未知类近邻关系学习合理聚类簇, 在开放条件下具备强可分特征表示能力(见图5(d))。

为充分评估不同标注比例下各方法性能, 图6给出了所提方法与对比方法在50%、100%已知类样本标注比例下的性能。由图6可知: 结合Sinkhorn-Knnp算法的隐空间学习方法: TRSSL、TIDA以及PKOSSL, 在已知类标注比例增加、无标记数据中已知类样本占比降低时, 因在 $K+U$ 类全局均匀分配伪标签易将未知类误分为已知类, 导致未知类聚类精度相比表1显著下降; 基于相似度学习的ORCA与OpenLDN因最大熵约束仅要求无标记样本向不同簇聚合, 50%标注下未知类精度未下降, 但100%标注下仍无法避免; 最新对比方法LeGoGCD和AFGCD分别基于局部熵正则与目标

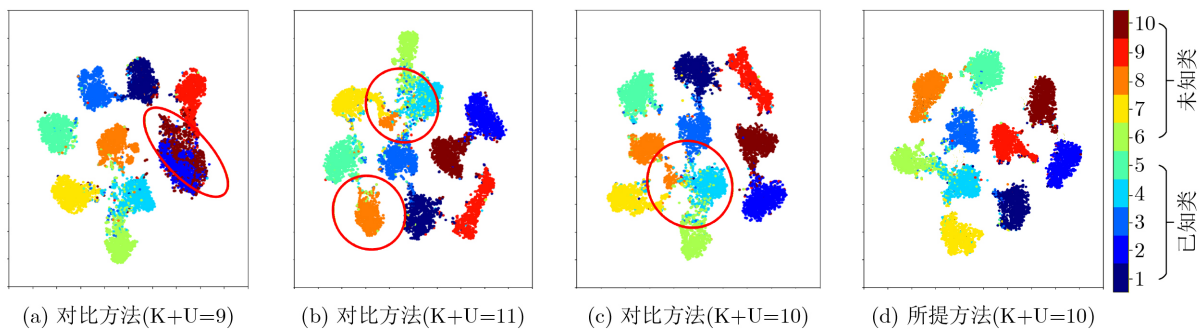


图5 所提方法与TRSSL(对比方法中整体性能最优)的无标记已知类/未知类特征t-SNE

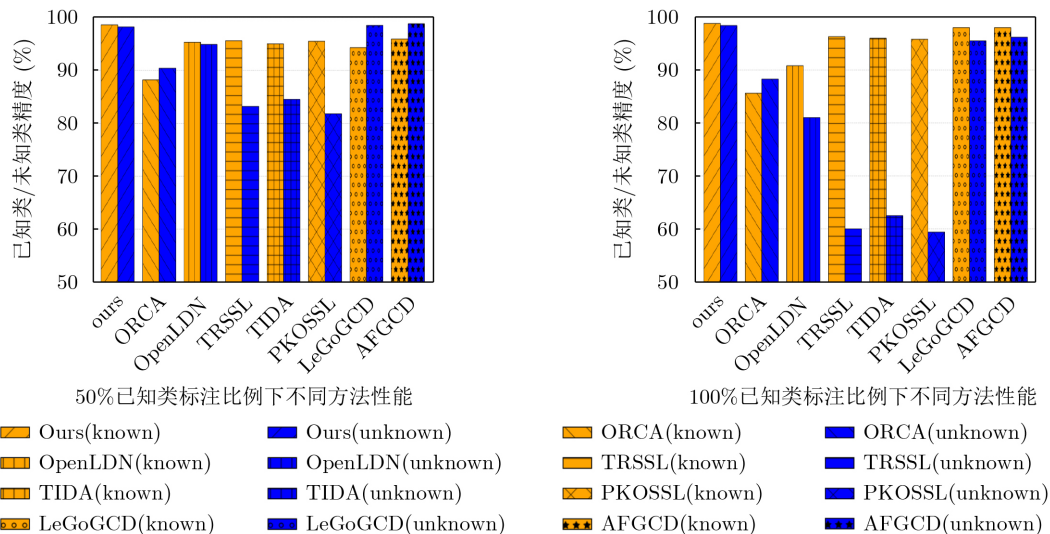


图6 所提方法与对比方法在50%、100%已知类标注比例下的已知类识别/未知类聚类精度(未知类个数已知条件下)

区域注意力聚焦增强已知-未知类间可分性，有效缓解未知类错判为已知类而精度下降的问题；所提方法显式建模已知-未知类鉴别边界(随标注比例提升建模更准确)，并为所鉴别已知、未知类分配互斥伪标签增强二者可分性，50%、100%标注比例下均最优。

(2)SAR图像数据集：MSTAR、ATRNet-STAR对比实验

SAR目标识别属典型开放场景(“非合作”特性导致未知新目标随机出现)，且SAR图像解译难度高于光学图像，对开放世界半监督学习需求迫切，但SAR领域尚无相关研究。本文在MSTAR、ATRNet-STAR数据集上对比所提方法与计算机视觉领域对比方法及无监督聚类下界(见表2)。特别地，SAR目标方位敏感性导致类内差异大、类间混淆，加剧OWSSL中未知类表征偏向已知类的问题。所提方法通过引入无标记已知类/未知类鉴别模块，约束已知、未知类样本向不同方向优化以减少错判。在MSTAR数据集上RAUS等指标分别优于对比方法5.69%、4.45%、5.76%以上，在类别繁

杂、已知-未知类更易混淆的ATRNet-STAR数据集上优势更显著。特征可视化(图7)验证所学已知/未知类可分性显著优于对比方法。

3.3.2 模型分析

本节从消融实验、无标记已知类-未知类模块分析、未知类自动聚类模块分析、模型超参数分析四方面对所提方法进行性能评估与合理性验证。

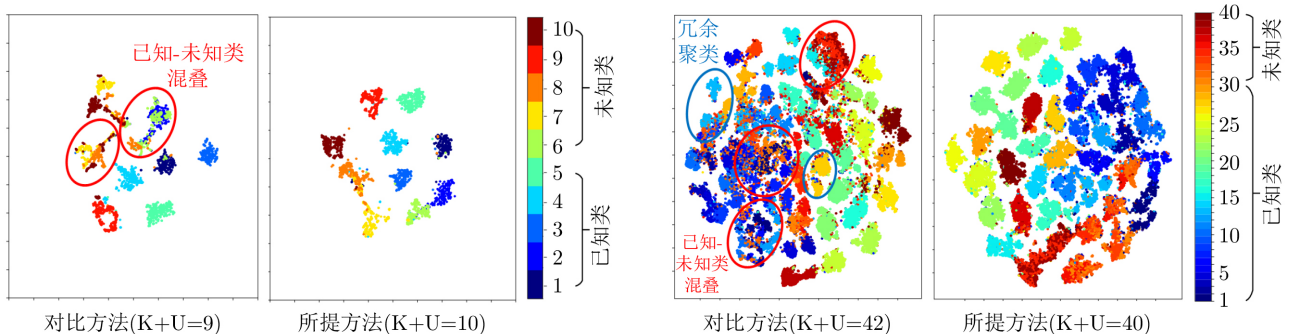
(1)消融实验

表3中基线方法(Baseline)为仅基于sinkhorn-knnp算法为无标记数据生成K+U类伪标签的方法。去除未知类自动聚类模块时，采用3.3.1节对比实验中介绍的K-means搜索方法确立K+U聚类个数。去除无标记已知类-未知类鉴别模块时，按前向推断结果对判定为未知类的无标记样本进行自动聚类。

由表3实验结果可知，本文所提无标记已知类-未知类鉴别模块和未知类自动聚类模块均能提升模型性能，二者组合时效果最优，验证了各模块有效性。光学数据CIFAR-10上，与基线相比，加入鉴别模块使得未知类拒判率RAUS提升7.84%，减少

表 2 不同方法在MSTAR、ATRNet-STAR数据集上的RAUS(%), ACC_{known}(%), ACC_{cluster}(%)(每个指标下左侧、右侧分别对应MSTAR数据、ATRNet-STAR数据实验，均在未知类个数未知条件下以三次K-means随机初始化后学习的平均结果作最终结果)及K+U估计个数

方法	RAUS		ACC _{known}		ACC _{cluster}		K+U		
	MSTAR	ATRNet-STAR	MSTAR	ATRNet-STAR	MSTAR	ATRNet-STAR	MSTAR	ATRNet-STAR	
无监督聚类	BootSC	-	72.80	68.62	61.18	25.99	[7,8,8]	[36,37,38]	
	SNSCC	-	75.15	70.33	62.02	26.96	[7,8,8]	[36,37,38]	
	ORCA	70.03	54.68	80.73	83.76	67.99	31.70	[8,9,9]	[38,39,41]
开放世界半监督学习	OpenLDN	72.32	58.53	83.09	88.12	70.66	34.24	[8,9,9]	[38,39,41]
	TRSSL	73.55	60.12	83.82	87.05	70.24	37.98	[8,9,9]	[38,39,41]
	PKOSSL	74.18	63.66	81.87	85.20	72.61	38.13	[8,9,9]	[38,39,41]
	LeGoGCD	80.90	75.08	60.95	89.17	78.45	42.22	[8,9,10]	[38,40,42]
	AFGCD	78.60	72.66	62.45	88.41	76.68	40.37	[8,9,10]	[38,40,42]
本文方法	86.59	84.94	88.27	90.15	84.21	50.09	[10,10,10]	[39,39,40]	



(a) 所提方法与对比方法可视化(结果来自MSTAR)

(b) 所提方法与对比方法可视化(结果来自ATRNet-STAR)

图 7 所提方法与对比方法(MSTAR、ATRNet-STAR对比中分别选择性能最优的TRSSL、LeGoGCD)的无标记已知类/未知类特征t-SNE

表 3 所提模块的消融实验结果RAUS、 ACC_{known} 、 ACC_{cluster} (每个指标下左侧、右侧分别对应CIFAR-10数据、ATRNet-STAR数据结果)

		无标记已知类-未知类鉴别模块	未知类自动聚类模块	RAUS		ACC_{known}		ACC_{cluster}	
Baseline	×	×		87.25	60.12	89.07	87.05	83.67	37.98
所提 模块	Π	×		95.09	80.97	91.54	88.60	86.35	44.85
	×	Π		86.76	62.79	90.19	87.33	89.01	39.01
	Π	Π		95.33	84.94	93.86	90.15	92.74	50.09

已知类-未知类间的错判(ACC_{known} 和 ACC_{cluster} 分别提升2.47%, 2.68%),且在类别繁杂、已知-未知类间易混淆的SAR数据ATRNet-STAR上提升更明显,但二者实验中未知类聚类性能 ACC_{cluster} 均受制于K-means搜索聚类个数与实际类别个数的偏差;仅加入聚类模块时,受未知类与已知类混淆影响, ACC_{cluster} 提升有限(CIFAR-10上 ACC_{cluster} 相比基线提升5.34%,ATRNet-STAR上仅提升1.03%);在鉴别模块基础上加入聚类模块使得 ACC_{cluster} 性能取得明显提升。

(2)基于已知类边界极值分布动态扩展的无标记已知类-未知类鉴别模块分析实验

图8(a)-(b)以某已知类为例展示极值距离拟合情况:仅用本类标记样本拟合时,其极值距离分布直方图、概率密度函数(Probability Density Function, PDF)及累积分布函数(Cumulative Distribution Function, CDF)曲线与本类全部样本差异显著,表明少量标记样本难以完整描述已知类边界(见图8(a));加入高置信样本后拟合结果与全部样本相近(见图8(b)),表明所提方法在半监督背景下结合高置信样本动态扩展已知类边界以表征完整边界分布的合理性,为无标记已知类-未知类鉴别提供可靠依据;该策略使无标记样本得以快速向不同方向优化,训练中期逼近最优收敛性能,而缺失该模块时已知、未知类在训练前期大量错判,且最终性能始终受制于错判样本(见图8(c))。图8(d)-(f)通过训练前、中、后期的决策分数(已知类特征与本类原型的余弦距离/未知类特征与最近已知类原型的余弦距离)统计直方图直观验证上述结论。图8(g)-(i)进一步展示训练前、中、后期的极值概率(以某已知类 k 极值分布为例,将本类特征与全部未知类特征-原型 p_k 距离代入极值分布)统计直方图,鉴别模块前期仅挑选高置信样本参与训练,并随无标记已知-未知类可分性提升、已知类分布趋于完善而动态调整鉴别阈值,最终实现所有无标记样本属于已知/未知类的准确判定。

(3)基于近邻交并比关系合并的未知类自动聚类模块分析实验

本方法以预设 $U_0 = 20$ 的K-means对无标记未

知类进行聚类初始化,并依据未知类各聚类簇近邻关系对相近冗余聚类簇进行合并,最终实现未知类合理聚类数 U^* 的自动确立。图9(a)-(d)给出了上述过程的特征可视化示意图:标记已知类预热训练后,基于 $U_0 = 20$ 的K-means进行聚类初始化并开始OWSSL学习,各未知类类内散布大、类间混淆明显;随已知类聚合知识为未知类聚类提供辅助,各未知类类内散布逐渐紧凑、类间分布逐渐推远,且同类近邻冗余聚类簇不断合并(如图9(b)-(c)所示),待自动聚类学习后基于确立的合理聚类数 U^* 优化至收敛,学得所有已知类、未知类类内紧凑、类间可分的特征(如图9(d)所示)。

(4)模型超参数分析实验

本文分别在光学数据CIFAR-10与SAR数据ATRNet-STAR上进行超参数分析实验(见图10),光学数据与SAR数据上超参数变化的性能变化趋势大体一致,具体而言:(a)当 $\tau_1 = 0.1$,大量无标记已知类无法参与训练使得 ACC_{known} 取得最低值, ACC_{cluster} 也较低。随 τ_1 增大 ACC_{known} 与 ACC_{cluster} 均上升。当 τ_1 过大($\tau_1 = 0.9$)时,大量无标记未知类被错判为已知类加入训练, ACC_{known} 与 ACC_{cluster} 均降低。所提方法通过动态确立 τ_1 ,在训练初期仅挑选高置信样本参与学习,并随高置信已知类动态扩展分布渐进提升 τ_1 ,最终纳入所有无标记样本, ACC_{known} 与 ACC_{cluster} 均取得最优;特别地,SAR数据ATRNet-STAR相比光学数据已知-未知类更易混淆,因而 ACC_{known} 与 ACC_{cluster} 均对 τ_1 变化更敏感;(b)当 $\tau_c = 0.5$,交并比不高的异类聚类簇被错误聚为一类, ACC_{cluster} 较低。随 τ_c 增大 ACC_{cluster} 上升。当 τ_c 过大($\tau_c = 1.0$)时,交并比趋近于1的聚类簇才能被合并,存在大量冗余簇, ACC_{cluster} 取得最低值。所提方法在自动聚类中自动搜索对已知类、未知类样本全局可分性最优的阈值 τ_c , ACC_{cluster} 取得最优(τ_c 在学习过程中动态变化,图10(b)仅展示了光学数据、SAR数据上自动聚类完成最终轮次搜索的阈值 τ_c 及其收敛性能);(c)当初始化聚类个数 U_0 小于实际未知类数 U^* (ATRNet-STAR中 $U_0 = 5, U^* = 10$)异类未知类簇将被强制合并, ACC_{cluster} 明显较低,当初始化聚类个数本身与实际

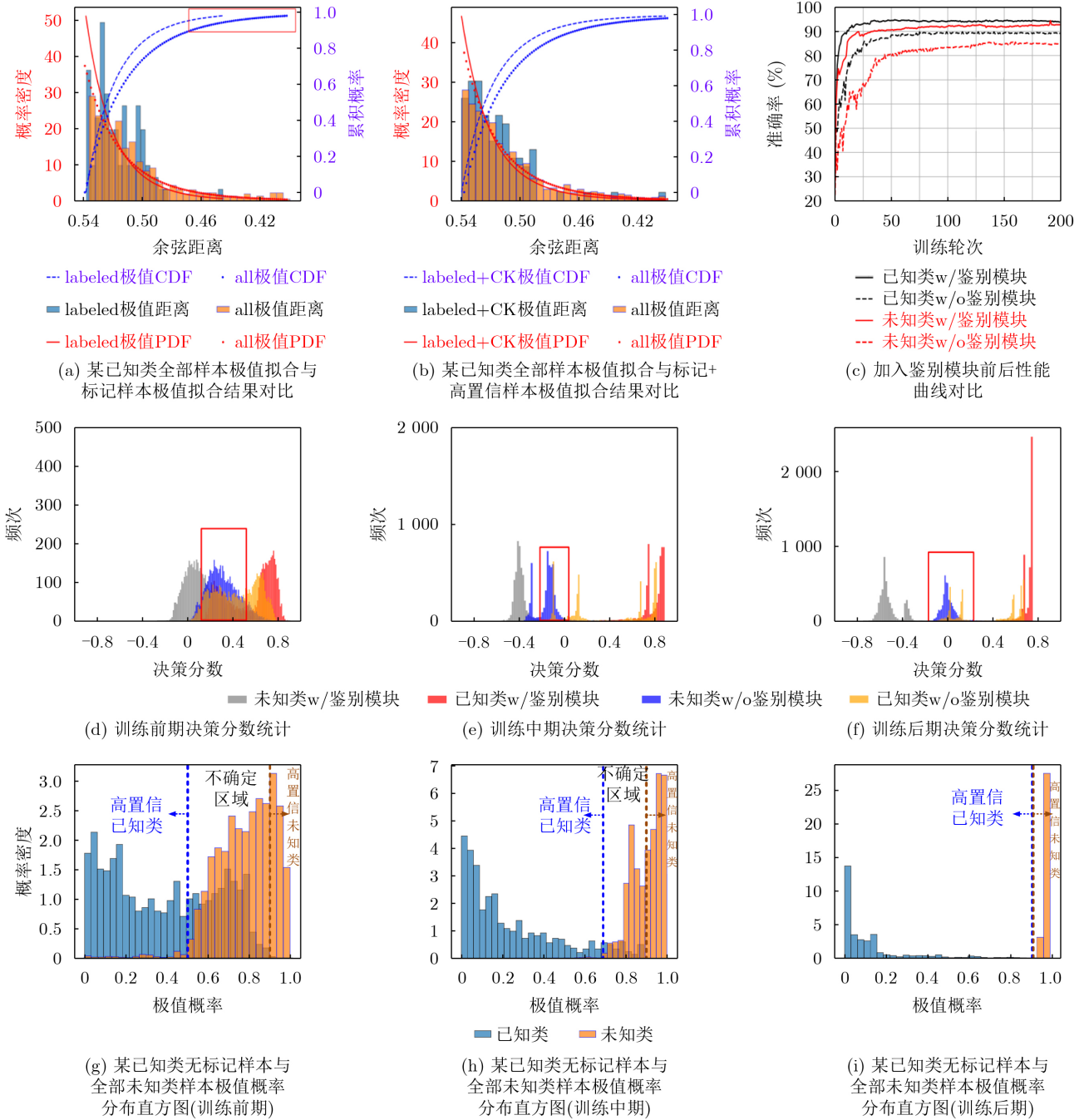


图 8 所提无标记已知类-未知类鉴别模块有效性验证示意图(结果来自CIFAR-10数据)

未知类类别数接近(CIFAR-10中 $U_0 = U^* = 5$, AT-RNet-STAR中 $U_0 = U^* = 10$), 若再执行所提自动合并聚类算法, 易导致部分异类未知类被错误聚为同一类, $ACC_{cluster}$ 相对较低。当 $U_0 \in [20, 40]$ 两数据集未知类聚类性能 $ACC_{cluster}$ 均具有稳健性, 因此应用中设置 U_0 为合理的较大值即可。

4 结论

针对开放条件下存在训练类别分布外的未知新目标亟待解译, 而传统闭集目标识别方法仅局限于

识别已知训练类别的问题, 本文研究了开放世界半监督学习方法, 采用直推式学习模式综合利用少量标记已知类数据、大量无标记已知类/未知类待测数据, 提升模型对已知类识别能力的同时实现对无标记未知新目标的发现与聚类。提出方法通过动态建模已知类边界分布实现无标记已知类/未知类鉴别, 并进一步基于未知类聚类簇的近邻关系实现未知类别数的自动估计与归属划分, 克服现有开放世界半监督学习方法缺乏显式已知/未知类鉴别策略致使未知类易错判, 以及局限于“未知类个数先验

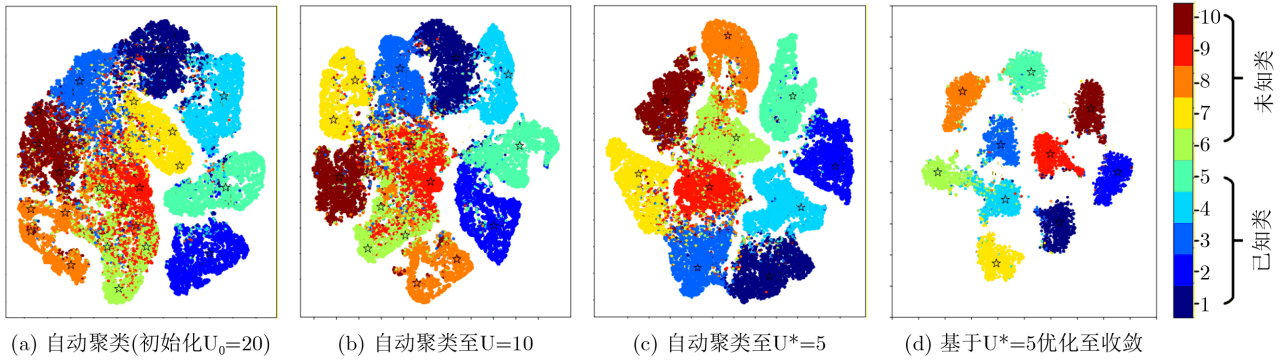


图 9 所提未知类自动聚类模块作用下的特征可视化示意图(其中星型表示类原型)(结果来自CIFAR-10数据)

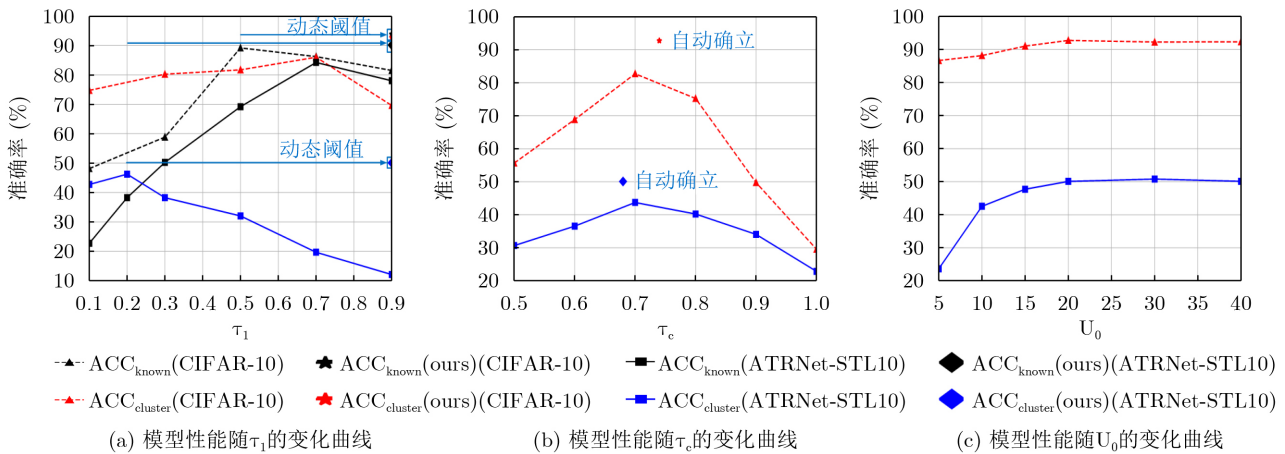


图 10 模型性能随超参数 τ_1 、 τ_c 、 U_0 的变化曲线

已知”强假设的问题。光学CIFAR-10数据和实测SAR数据实验结果表明,所提方法在开放条件下实现了已知类的准确识别与未知类的准确聚类。

参考文献

- [1] 张东阳, 陆子轩, 刘军民, 等. 深度模型的持续学习综述: 理论、方法和应用[J]. 电子与信息学报, 2024, 46(10): 3849–3878. doi: [10.11999/JEIT240095](https://doi.org/10.11999/JEIT240095).
ZHANG Dongyang, LU Zixuan, LIU Junmin, *et al.* A survey of continual learning with deep networks: Theory, method and application[J]. *Journal of Electronics & Information Technology*, 2024, 46(10): 3849–3878. doi: [10.11999/JEIT240095](https://doi.org/10.11999/JEIT240095).
- [2] . CAO Kaidi, BRBIC M, and LESKOVEC J. Open-world semi-supervised learning[C]. International Conference on Learning Representations, 2022: 22837–22855. (查阅网上资料, 未找到本条文献的出版地信息, 请确认).
- [3] . RIZVE M N, KARDAN N, KHAN S, *et al.* Openldn: Learning to discover novel classes for open-world semi-supervised learning[C]. 17th European Conference on Computer Vision–ECCV 2022, Tel Aviv, Israel, 2022: 382–401. doi: [10.1007/978-3-031-19821-2_22](https://doi.org/10.1007/978-3-031-19821-2_22).
- [4] . RIZVE M N, KARDAN N, and SHAH M. Towards realistic semi-supervised learning[C]. 17th European Conference on Computer Vision – ECCV 2022, Tel Aviv, Israel, 2022: 437–455. doi: [10.1007/978-3-031-19821-2_25](https://doi.org/10.1007/978-3-031-19821-2_25).
- [5] . MARCO C. Sinkhorn distances: Lightspeed computation of optimal transport[C]. Advances in Neural Information Processing Systems, Stateline 26: 27th Annual Conference on Neural Information Processing Systems 2013, Lake Tahoe, United States, 2013: 2292–2300.
- [6] ZHAO Tianhao, LIN Yutian, WU Yu, *et al.* Promote knowledge mining towards open-world semi-supervised learning[J]. *Pattern Recognition*, 2024, 149: 110259. doi: [10.1016/j.patcog.2024.110259](https://doi.org/10.1016/j.patcog.2024.110259).
- [7] . CAO Xinzi, ZHENG Xiawu, WANG Guanhong, *et al.* Solving the catastrophic forgetting problem in generalized category discovery[C]. Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, USA, 2024, 16880–16889. doi: [10.1109/CVPR52733.2024.01597](https://doi.org/10.1109/CVPR52733.2024.01597).
- [8] . XU Qiyu, HU Zhanxuan, DUAN Yu, *et al.* A hidden stumbling block in generalized category discovery: Distracted attention[C]. Proceedings of the 2025 IEEE/CVF International Conference on Computer Vision (ICCV), Honolulu, USA, 2025: 405–414. doi: [10.1109/ICCV51701](https://doi.org/10.1109/ICCV51701).

- 2025.00045.
- [9] . KRIZHEVSKY A and HINTON G E. Learning multiple layers of features from tiny images[R]. 2009. (查阅网上资料, 未找到本条文献的报告编号信息, 请确认).
- [10] ROSS T D, WORRELL S W, VELTEN V J, *et al.* Standard SAR ATR evaluation experiments using the MSTAR public release data set[J]. *Algorithms for Synthetic Aperture Radar Imagery V*, 1998, 3370: 566–573. doi: [10.1117/12.321859](https://doi.org/10.1117/12.321859).
- [11] LIU Yongxiang, LI Weijie, LIU Li, *et al.* ATRNet-STAR: A large dataset and benchmark towards remote sensing object recognition in the wild[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026, 48(6): 6735–6753. doi: [10.1109/TPAMI.2026.3658649](https://doi.org/10.1109/TPAMI.2026.3658649).
- [12] . MOVSHOVITZ-ATTIAS Y, TOSHEV A, LEUNG T K, *et al.* No fuss distance metric learning using proxies[C]. Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017: 360–368. doi: [10.1109/ICCV.2017.47](https://doi.org/10.1109/ICCV.2017.47).
- [13] 陈健, 雍奇锋, 杜兰, 等. 结合未知类特征生成与分类得分修正的SAR目标开集识别方法[J]. 电子与信息学报, 2024, 46(10): 3890–3907. doi: [10.11999/JEIT240138](https://doi.org/10.11999/JEIT240138).
CHEN Jian, YONG Qifeng, DU Lan, *et al.* An open set recognition method for SAR targets combining unknown feature generation and classification score modification[J]. *Journal of Electronics & Information Technology*, 2024, 46(10): 3890–3907. doi: [10.11999/JEIT240138](https://doi.org/10.11999/JEIT240138).
- [14] 杜兰, 李逸明, 薛世鲲, 等. 结合相似度预测和阈值自动求解的开集条件下毫米波雷达点云步态识别方法[J]. 电子与信息学报, 2025, 47(6): 1850–1863. doi: [10.11999/JEIT241034](https://doi.org/10.11999/JEIT241034).
DU Lan, LI Yiming, XUE Shikun, *et al.* Millimeter-wave radar point cloud gait recognition method under open-set conditions based on similarity prediction and automatic threshold estimation[J]. *Journal of Electronics & Information Technology*, 2025, 47(6): 1850–1863. doi: [10.11999/JEIT241034](https://doi.org/10.11999/JEIT241034).
- [15] GUO Wengang, YE Wei, CHEN Chunchun, *et al.* Bootstrap deep spectral clustering with optimal transport[J]. *IEEE Transactions on Multimedia*, 2026, 28: 531–544. doi: [10.1109/TMM.2025.3623492](https://doi.org/10.1109/TMM.2025.3623492).
- [16] DUAN Yu, CHEN Huimin, ZHANG Runxin, *et al.* Soft neighbors supported contrastive clustering[J]. *IEEE Transactions on Image Processing*, 2025, 34: 4315–4327. doi: [10.1109/TIP.2025.3583194](https://doi.org/10.1109/TIP.2025.3583194).
- 宋佳伦：男，博士生，研究方向为开放条件目标识别、增量学习、小样本学习。
杜 兰：女，博士，教授，博士生导师，主要研究方向为雷达目标识别、雷达信号处理、机器学习。
陈 健：男，博士，副教授，硕士生导师，主要研究方向为雷达目标检测识别与机器学习。

责任编辑：余 蓉

Explicit Discrimination-Driven Automatic Unknown Class Clustering for Open-World Semi-Supervised Learning

SONG Jialun^① DU Lan^① CHEN Jian^①

^①(National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China)

Abstract: Traditional target recognition follows closed-set assumption, restricting test targets to those present in training data, yet real-world scenarios involve unlabeled unknown classes requiring simultaneous recognition of known categories and discovery/clustering of novel ones. Thus, this paper introduces a transductive open-world semi-supervised learning method based on explicit discrimination-driven automatic unknown class clustering, leveraging limited labeled known-class data and extensive unlabeled testing samples (knowns and unknowns). The approach combines: Dynamic Known-Unknown Class Discrimination (DKUCD), which models labeled known class boundaries using extreme value theory and refines them dynamically with high-confident knowns samples to enhance discrimination, and Neighbor Intersection-over-Union Clustering (NIOUC), which clusters high-confidence unknowns samples by merging neighbors based on intersection-over-union ratios. DKUCD and NIOUC are optimized iteratively. Experiments on optical CIFAR-10 and measured radar data demonstrate the method's effectiveness in accurately recognizing known classes while clustering unknown classes.

Key words: Target recognition; Open-world semi-supervised learning; Unknown class discrimination; Automatic clustering; Extreme value theory.