

基于双频配对分段与多表示学习的无人机射频识别方法

林成琨, 郝东来*, 何天洋, 王加瑞

西京学院计算机学院, 陕西 西安

收稿日期: 2026年8月4日; 录用日期: 2026年9月4日; 发布日期: 2026年9月11日

摘 要

DroneRF以长序列形式存储射频数据, 先生成短窗口再随机划分可能使同一低频段/高频段(L/H)配对分段的相关窗口进入不同数据子集。本文将454个单频段CSV文件整理为227个L/H配对分段, 在窗口生成前划分为160、34和33个训练、验证与测试配对分段; 每个配对分段生成32个尺寸为 2×8192 的双频索引对齐窗口, 共计7264个窗口, 并固定数据清单、测试清单和指标规则。在该协议下, 采用Raw-FFT-STFT三支网络比较CE、MildAug、Source-aware SupCon及联合配置。三个随机种子下, 联合配置按窗口统计的平衡准确率为 0.9926 ± 0.0051 , 较CE的 0.9901 ± 0.0037 提高0.25个百分点, 但该变化未在全部随机种子中保持一致; 四种配置汇总33个测试配对分段后均正确分类。去除Raw和STFT分支后, 平衡准确率分别下降0.0059和0.0022; 去除FFT分支后的变化为+0.0006, 处于实验波动范围内。固定随机种子下, 非零对比损失权重均未超过CE; Source-aware SupCon提高了跨配对分段的同类相似度, 但伴随类别间隔收缩。所建协议为固定数据与统一规则下的方法比较提供了可复核基础, 结论限于DroneRF现有文件信息与Background/UAV二分类任务。

关键词

无人机射频识别, DroneRF, L/H配对分段, 防泄漏评价协议, 多表示学习, 监督对比学习

RF-Based UAV Recognition with Dual-Band Paired Segments and Multi-Representation Learning

Chengkun Lin, Donglai Hao*, Tianyang He, Jiarui Wang

School of Computer Science, Xijing University, Xi'an Shaanxi

Received: August 4, 2026; accepted: September 4, 2026; published: September 11, 2026

*通讯作者。

文章引用: 林成琨, 郝东来, 何天洋, 王加瑞. 基于双频配对分段与多表示学习的无人机射频识别方法[J]. 计算机科学与应用, 2026, 16(9): 140-153. DOI: 10.12677/csa.2026.169296

Abstract

DroneRF stores radio-frequency measurements as long sequences. Randomly splitting short windows may place correlated windows from the same low-/high-band (L/H) paired segment in different subsets and reduce evaluation reliability. This study therefore organizes 454 single-band CSV files into 227 L/H paired segments according to BUI activity codes and segment indices, and assigns 160, 34, and 33 paired segments to the training, validation, and test sets before window generation. Each paired segment yields 32 index-aligned dual-band windows of size 2×8192 , producing 7264 windows. Data manifests, the test list, and metric definitions are fixed, and identifier exclusivity is audited. Under this protocol, a Raw-FFT-STFT three-branch network is evaluated with cross-entropy (CE), mild data augmentation (MildAug), Source-aware SupCon, and their joint configuration. Across three random seeds, the joint configuration achieves a per-window balanced accuracy of 0.9926 ± 0.0051 versus 0.9901 ± 0.0037 for CE, an increase of 0.25 percentage points that is not retained for every seed. Probability aggregation correctly classifies all 33 test paired segments for all four configurations. Removing the Raw and STFT branches reduces balanced accuracy by 0.0059 and 0.0022, respectively, whereas removing FFT changes it by +0.0006, within experimental variation. With a fixed seed, no nonzero contrastive-loss weight exceeds CE. Source-aware SupCon increases within-class similarity across paired segments but reduces class separation. The protocol provides a reproducible basis for comparison under fixed data and evaluation rules; the conclusions are limited to the available DroneRF metadata and the Background/UAV binary task.

Keywords

RF-Based UAV Recognition, DroneRF Dataset, L/H Paired Segments, Leakage-Resistant Evaluation Protocol, Multi-Representation Learning, Supervised Contrastive Learning

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,无人机(Unmanned Aerial Vehicle, UAV)广泛应用于航空测绘、物流运输、农业监测和灾害救援等领域。与此同时,未经授权的无人机活动也给机场、重要设施及大型活动场所带来了低空安全风险。因此,实现无人机目标的快速、准确识别已成为低空安全感知领域的重要研究内容[1]。

视觉、雷达、声学 and 射频(Radio Frequency, RF)均可用于无人机感知。其中,射频感知通过接收无人机与遥控器之间的控制、遥测或图传信号进行判别,具有被动接收、不依赖目标可见性等特点。随着深度学习的发展,卷积神经网络等模型开始用于处理原始射频率序列及其频域、时频表示,降低了对人工特征设计的依赖[2]。

Al-Sa'd 等基于深度学习开展无人机检测、型号识别和运行状态分类研究[3], Allahham 等公开了包含背景信号及三种无人机双频段射频信号的 DroneRF 数据集,并说明了数据采集对象、工作状态和文件组织方式[4]; Nemer 等进一步采用层次化学习框架完成无人机存在性检测、型号识别和运行状态分类[5]。上述研究验证了射频信号和数据驱动方法在无人机识别任务中的应用可行性。

DroneRF 原始数据以长时间连续信号形式存储,现有文件可依据活动代码、频段标记和分段编号恢

复为低频/高频(L/H)配对分段。若先将连续信号切分为大量短窗口，再以窗口为单位随机划分训练集和测试集，同一配对分段中的相关窗口可能跨数据子集分布。模型由此可能利用特定分段的设备状态、传播环境、背景噪声或局部信号结构完成判别，使测试结果不能充分反映其对未参与训练分段的识别能力。针对无人机射频基准的控制研究表明，切片后随机划分可能显著高估模型性能[6]；机器学习研究中的一般性证据也指出，未按数据来源分组的划分方式是造成评价偏差和复现风险的重要因素[7]。

针对上述问题，本文构建面向 DroneRF 数据集的 L/H 配对分段互斥评价协议。首先，根据共同的 BUI 活动代码和分段编号，将 454 个单频段 CSV 文件整理为 227 个 L/H 配对分段；随后在窗口生成之前，按配对分段完成训练集、验证集和测试集的互斥划分；最后从每个配对分段构建 32 个长度为 8192 点的双频索引对齐窗口，共形成 7264 个窗口。训练过程使用类别 - 配对分段均衡采样器，验证集和测试集使用固定 manifest 及固定清单，以保证不同训练配置在相同条件下进行比较。

在该协议下，本文以 Raw-FFT-STFT 多表示分类网络为统一实验骨干，比较交叉熵、轻量射频增强 (MildAug)、配对分段感知监督对比学习 (Source-aware SupCon) 及其联合配置。Source-aware SupCon 以监督对比学习[8]为基础，仅将类别相同且来自不同 L/H 配对分段的样本构造为有效正样本，以探索跨分段同类特征约束的作用。

本文的主要工作和贡献如下。

(1) 构建 DroneRF 双频数据的 L/H 配对分段互斥评价协议，在窗口生成前完成数据集划分，从数据流程上阻断同一配对分段窗口跨训练集、验证集和测试集分布的直接泄漏路径。

(2) 建立固定 manifest、固定测试清单和统一评价规则，并对单频段 CSV 路径、L/H 配对分段标识及窗口标识进行互斥性审计，使不同训练配置的输入数据、测试对象和指标计算过程保持一致。

(3) 在上述协议下开展 Raw、FFT 与 STFT 分支消融，以及 MildAug 和 Source-aware SupCon 的受控实验，并结合损失权重与温度参数敏感性和特征空间分析，客观报告不同训练组件的混合结果。

三个随机种子下，各训练配置均取得较高的按窗口统计识别性能；联合配置的平均平衡准确率略高于交叉熵对照方法，但未在全部随机种子中保持优势。消融与参数实验表明，Raw 分支的作用相对明确，Source-aware SupCon 对跨分段同类聚集和类别间隔的影响并不一致。因此，本文将其作为严格评价协议下的探索性训练配置，不概括为稳定的性能改进方法。

本文其余内容安排如下：第二章综述相关研究；第三章介绍配对分段互斥协议、多表示网络与训练方法；第四章给出对照、消融、参数敏感性和特征空间实验；第五章总结研究结论、局限与后续方向。

2. 相关研究进展

2.1. 基于射频信号的无人机识别研究

无人机飞行产生的控制、遥测和图传信号可用于目标检测与识别。射频感知采用被动接收方式，不依赖目标可见性，可作为视觉、雷达和声学感知的补充[1]。

早期方法通常从时域波形、功率谱或统计量中提取人工特征，再利用支持向量机、K 近邻等分类器完成判别，其性能易受特征设计和采集条件变化影响。随着深度学习的发展，神经网络能够从原始射频序列及其频域表示中自动学习特征[2] [9]；Al-Sa'd 等据此实现了无人机检测、型号识别和运行状态分类[3]。

现有研究证明了深度模型的有效性，但模型性能还受窗口构建和数据划分方式影响。对于由少量连续分段生成的大量窗口，若划分时忽略分段归属，模型可能利用特定采集来源的局部模式，导致测试结果偏高[6] [7]。因此，评价射频识别方法时应同时说明窗口生成、划分单位与测试对象。

2.2. DroneRF 数据集及识别方法研究

DroneRF 是具有代表性的公开无人机射频数据集,相关工作分别开展了无人机检测、型号识别和运行状态分类[3],并系统说明了数据采集对象、工作状态与文件组织方式[4]。该数据集包含背景信号及 Parrot AR、Parrot Bebop 和 DJI Phantom 三种无人机在不同运行状态下的双频段数据,为相关方法比较提供了公开基础。

基于 DroneRF, Nemer 等构建层次化识别框架[5], Mo 等将压缩感知与深度网络结合以降低射频数据规模[10]。相关工作推动了 Raw、FFT 和 STFT 等多种信号表示在无人机识别中的应用。

DroneRF 由有限数量的长时间连续分段构成,模型输入通常由其进一步切片得到。若切片后随机划分,同一分段的相关窗口可能同时进入训练集和测试集,导致对未参与训练分段识别能力的高估[6];相关研究亦将此类相关数据跨集合分布视为影响可复现性的重要泄漏来源[7]。因此,本文在窗口生成前执行 L/H 配对分段互斥划分。由于公开数据缺少独立采集过程标识,L/H 配对分段仅作为可审计的划分单位,不等同于完全独立地采集会话。

2.3. 射频信号表示与深度特征学习研究

射频信号可从时域、频域和时频域表征。Raw 序列保留幅值变化与局部结构,卷积网络可直接学习其判别特征[2] [9];FFT 描述窗口内频率分量的整体分布,STFT 进一步刻画频谱随时间的变化[11]。三类表示信息互补,但其实际贡献仍需在网络和评价条件下验证。

本文采用 Raw-FFT-STFT 三支网络作为统一分类骨干,分别编码原始序列、幅度谱和时频图,融合后完成 Background/UAV 二分类。所有训练配置使用相同网络,各分支作用通过逐分支消融实验评价。

2.4. 数据泄漏控制与监督对比学习

连续信号切分后,同一数据来源的窗口通常共享设备状态、传播环境、背景噪声和局部结构。控制直接泄漏的基本条件是保留来源标识,并使同一来源的全部窗口归入同一数据子集[6] [7]。但分组划分只能保证现有标识单元互斥,不能据此推断设备、环境或独立采集过程完全独立。

监督对比学习(Supervised Contrastive Learning, SupCon)依据类别标签构造正样本关系,直接约束批次内特征分布[8]。标准 SupCon 不区分同类窗口的分段来源,可能强化分段特有模式。本文将 L/H 配对分段标识引入正样本掩码,仅把类别相同且来自不同配对分段的窗口定义为有效正样本;其中“Source”指由文件组织关系恢复的配对分段来源,不表示领域适应中的源域。

领域对抗神经网络(DANN)通过域分类器和梯度反转层抑制域可辨识信息[12],Group DRO 则针对预定义群组优化最坏群组风险[13]。两者分别依赖可信的域标签或具有实际意义的群组定义。DroneRF 现有配对分段标识不能证明独立设备、场景或采集会话,因而本文仅将 DANN/GRL 与 Group DRO 作为概念比较对象,并将 Source-aware SupCon 的结论限定于现有标识下的跨分段特征约束。

3. 基于双频配对分段与多表示学习的 DroneRF 射频识别方法

3.1. 问题定义与整体框架

DroneRF 原始射频数据以长序列保存,同一低频段/高频段(L/H)配对分段可生成多个相关窗口。若先生成窗口再随机划分,同一配对分段的局部信号模式可能同时进入训练集与测试集,影响评价结果。为控制这一问题,本文将 L/H 配对分段作为数据组织、划分、采样和结果汇总单位,并在窗口生成前固定训练集、验证集和测试集。该单位依据文件名恢复,仅表示可审计的数据来源关系,不等同于独立设备或独立采集过程。

本文方法包括配对分段互斥划分、双频索引对齐窗口构建、多表示特征学习和受控训练四个环节。454 个单频段 CSV 文件首先整理为 227 个 L/H 配对分段，每个分段生成 32 个尺寸为 2×8192 的窗口；随后由 Raw、FFT 和 STFT 三支网络提取特征，并通过类别——配对分段均衡采样、轻量射频增强 (MildAug) 和配对分段感知监督对比学习开展训练。整体流程见图 1。

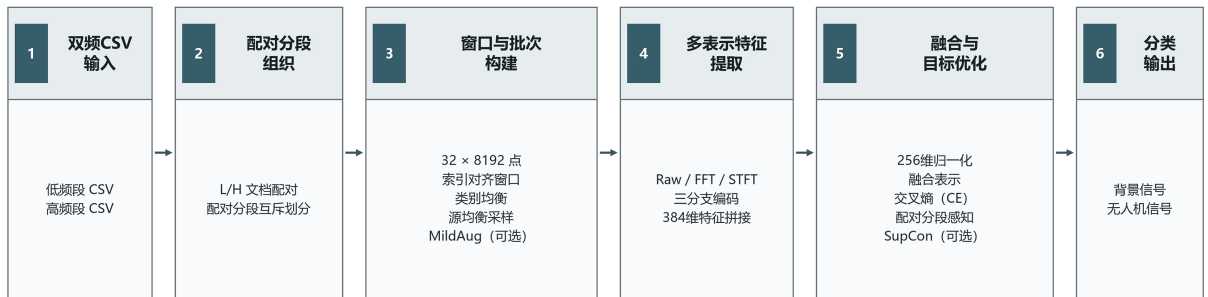


Figure 1. Overall procedure of RF signal classification based on L/H paired segments
图 1. 基于 L/H 配对分段的射频信号分类整体流程

3.2. L/H 配对分段构建与窗口采样

根据 DroneRF 文件命名规则，具有相同 BUI 活动代码和分段编号的低频段文件与高频段文件组成一个 L/H 配对分段。454 个单频段 CSV 文件共形成 227 个配对分段，其中 Background 为 41 个，AR、Bebop 和 Phantom 合并后的 UAV 为 186 个。训练集、验证集和测试集分别分配 160、34 和 33 个配对分段。设三者的配对分段标识集合为 G_{tr} 、 G_{va} 和 G_{te} ，则划分满足：

$$G_{tr} \cap G_{va} = \emptyset, G_{tr} \cap G_{te} = \emptyset, G_{va} \cap G_{te} = \emptyset \quad (1)$$

同一配对分段的低频段文件、高频段文件及其生成的全部窗口始终归入同一数据子集。固定数据清单进一步核对 CSV 路径、频段文件标识、配对分段标识和窗口标识，未发现跨训练集、验证集和测试集的直接重复。各数据子集的构成见表 1。

Table 1. Split results of L/H paired segments and windows
表 1. 配对分段及窗口划分结果

数据集	Background 分段	UAV 分段	分段总数	Background 窗口	UAV 窗口	窗口总数
训练集	29	131	160	928	4192	5120
验证集	6	28	34	192	896	1088
测试集	6	27	33	192	864	1056
合计	41	186	227	1312	5952	7264

每个单频段 CSV 文件包含 10,000,000 个采样点。为覆盖完整信号分段，本文将序列划分为 32 个等宽区间，并由固定随机种子 3407 与配对分段标识共同确定各区间的窗口中心。数据构建阶段保留 8192、16,384 和 32,768 三种候选长度，并按最大候选长度预留边界；正式实验固定采用 $\ell = 8192$ 。低频段与高频段使用相同的窗口编号、中心位置和数值索引，双频输入表示为：

$$\mathbf{x}_{r,j}^{(\ell)} = \begin{bmatrix} \mathbf{s}_r^L \left[c_{r,j} - \frac{\ell}{2} : c_{r,j} + \frac{\ell}{2} \right) \\ \mathbf{s}_r^H \left[c_{r,j} - \frac{\ell}{2} : c_{r,j} + \frac{\ell}{2} \right) \end{bmatrix} \in \mathbb{R}^{2 \times \ell} \quad (2)$$

每个 L/H 配对分段生成 32 个互不重叠的双频窗口，最终得到 7264 个尺寸为 2×8192 、数据类型为 float32 的样本。索引对齐保证两路文件在数据构建过程中的位置对应，但公开数据未提供硬件时钟信息，因而不将其解释为硬件时钟级绝对同步。现有文件信息也不能证明不同配对分段来自完全独立的设备、环境或采集过程。

3.3. Raw-FFT-STFT 多表示分类网络

三种表示均由同一个双频时域窗口生成。模型首先按样本、按频段通道沿时间维进行窗口内部标准化，不使用训练集、验证集或测试集的全局统计量。Raw 分支直接接收标准化时域信号；FFT 分支对单边实数频谱的幅值进行对数压缩和标准化：

$$X_{\text{fft}} = \text{Norm} \left[\log \left(1 + \left| \text{rFFT} (X_{\text{raw}}) \right| \right) \right] \quad (3)$$

STFT 分支采用长度为 256 的 Hann 窗，傅里叶变换长度为 256，滑动步长为 128，并对幅度时频谱进行对数压缩和样本内部标准化：

$$X_{\text{stft}} = \text{Norm} \left[\log \left(1 + \left| \text{STFT} (X_{\text{raw}}; n_{\text{fft}} = 256, w = 256, h = 128) \right| \right) \right] \quad (4)$$

Raw 和 FFT 分支分别使用参数独立的五层一维卷积编码器，通道数依次为 32、64、128、192 和 256，卷积核尺寸依次为 15、11、7、5 和 3，步长均为 2。STFT 分支采用五层二维卷积编码器，通道数为 24、48、96、128 和 192，卷积核均为 3×3 ，首层步长为 1，其余层步长为 2。各卷积块均由卷积、批归一化、GELU 和 Dropout 组成；分支末端结合自适应平均池化与最大池化，分别输出 128 维特征。三路特征拼接为：

$$\mathbf{f}_{\text{cat}} = \text{Concat}(\mathbf{f}_{\text{raw}}, \mathbf{f}_{\text{fft}}, \mathbf{f}_{\text{stft}}) \in \mathbb{R}^{384} \quad (5)$$

384 维拼接特征依次经过线性映射、LayerNorm、GELU、Dropout 和第二个线性映射，形成 256 维融合特征，并经 L2 归一化后同时用于线性二分类器和监督对比损失。完整三支网络包含 1,440,993 个可训练参数。第四章在相同数据清单和训练条件下进行逐分支消融；由于去除分支会同时改变网络结构与参数量，相关结果仅用于评价当前结构中的边际变化。

3.4. 类别 - 配对分段均衡采样与轻量增强

训练集包含 928 个 Background 窗口和 4192 个 UAV 窗口。为降低类别数量差异及少数配对分段重复出现的影响，批量大小固定为 32，每个批次等量抽取 16 个 Background 窗口和 16 个 UAV 窗口，并在类别内部轮换选择不同 L/H 配对分段及其窗口。每轮训练包含 160 个批次，共训练 100 轮。该过程允许数量较少的 Background 窗口重复参与训练，并非对 5120 个训练窗口进行一次无放回遍历；全部实验均采用普通交叉熵损失，不额外设置类别权重。

MildAug 由弱高斯噪声和有限时间平移组成。对第 i 个样本的第 c 个频段通道，噪声标准差设为当前通道标准差的 0.005 倍；整数平移量 Δ_i 从 $[-64, 64]$ 均匀抽取，并由同一样本的低频段和高频段共享。增强后的信号为：

$$\tilde{\mathbf{x}}_{i,c} = \text{Roll}_{\Delta_i}(\mathbf{x}_{i,c} + \epsilon_{i,c}) \quad (6)$$

时间平移采用循环回卷方式，不使用补零或反射填充。MildAug 仅在训练阶段启用，验证集和测试集使用固定原始窗口；增强先作用于双频时域窗口，随后由同一增强结果生成 Raw、FFT 和 STFT 表示，从而保证三个分支接受一致的随机扰动。

3.5. 配对分段感知监督对比学习

标准监督对比学习依据类别标签构造正样本关系[8]。对于由连续分段切分得到的窗口，仅按类别选择正样本可能反复利用同一配对分段内部的相关模式。本文在正样本掩码中加入 L/H 配对分段标识。设训练批次索引集合为 B ，第 i 个样本的类别标签、配对分段标识和 L2 归一化特征分别为 y_i 、 g_i 和 z_i ，则有效正样本集合定义为：

$$P(i) = \{p \in B \setminus \{i\} | y_p = y_i, g_p \neq g_i\} \quad (7)$$

只有类别相同且来自不同 L/H 配对分段的样本进入正样本集合。同类别、同配对分段样本仍保留在对比损失分母中；仅对 $P(i)$ 非空的锚点计算损失，记有效锚点集合为 $I^+ = \{i \in B | |P(i)| > 0\}$ 。配对分段感知监督对比损失为：

$$\mathcal{L}_{ps} = -\frac{1}{|I^+|} \sum_{i \in I^+} \frac{1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(z_i^T z_p / \tau)}{\sum_{a \in B \setminus \{i\}} \exp(z_i^T z_a / \tau)} \quad (8)$$

温度参数 τ 设为 0.1。模型不设置独立投影头，也不构造双增强视图；对比损失直接作用于与分类器共享的 256 维归一化特征。启用该训练方式时，总损失由普通交叉熵损失与配对分段感知监督对比损失共同构成：

$$L = \mathcal{L}_{ce} + \lambda_{ps} \mathcal{L}_{ps} \quad (9)$$

监督对比损失权重 λ_{ps} 设为 0.03，训练批次中每个锚点均具有 14~15 个有效正样本。第四章沿用实验名称“Source-aware SupCon”，其中“Source”仅指由文件组织关系恢复的 L/H 配对分段来源，不表示领域适应中的源域。该方法约束的是现有配对分段之间的特征关系，不据此推断跨设备或跨环境泛化能力。

4. 实验结果与分析

4.1. 实验设置与协议核对

实验在 Windows 11、AMD Ryzen 7 9700X、32 GB 内存和 NVIDIA GeForce RTX 5070 12 GB 平台上完成，程序采用 Python 3.13、PyTorch 2.10.0 和 CUDA 实现。模型输入尺寸为 2×8192 的双频索引对齐窗口，Raw、FFT 和 STFT 分支各输出 128 维特征，经融合模块得到 256 维归一化表示；完整网络包含 1,440,993 个可训练参数。全部实验使用 AdamW 优化器，初始学习率与权重衰减均为 5×10^{-4} ，批量大小为 32，每轮 160 个批次，训练 100 轮。

训练批次等量抽取 16 个 Background 窗口和 16 个 UAV 窗口，并在类别内部轮换不同 L/H 配对分段。四种训练配置分别为 CE、CE + MildAug、CE + Source-aware SupCon 及两者联合；MildAug 采用标准差系数 0.005 的弱高斯噪声和 ± 64 点共享时间平移，监督对比损失权重 $\lambda_{ps} = 0.03$ 、温度 $\tau = 0.1$ ，其余条件一致。每种配置使用随机种子 42、2024 和 3407 独立训练。模型依据固定验证集的配对分段汇总 RS、按窗口统计的 RS 及交叉熵损失依次选择，测试集不参与模型选择。

评价指标包括 Accuracy、Balanced Accuracy (BA)、两类召回率、F1、Minimum Recall、Recall Gap 和 Reliability Score (RS)。考虑类别数量差异，方法比较主要依据按窗口统计的 BA、召回率、F1 与错误分布；同一 L/H 配对分段内 32 个窗口的平均概率用于补充验证。固定清单的数量及互斥性核对见表 2。

三个集合在窗口生成前分别分配 160、34 和 33 个 L/H 配对分段，同一配对分段产生的 32 个窗口始终保留在同一集合。CSV 路径、频段文件标识、配对分段标识和窗口标识均未发现跨集合直接重复。DroneRF 未提供设备编号、采集日期和独立采集过程标识，因此该核对仅验证现有清单字段下的互斥性，不用于推断跨设备或跨环境泛化能力。

Table 2. Verification of the fixed data lists and split
表 2. 固定数据清单与划分检查

检查项目	训练集	验证集	测试集	合计与结果
单频 CSV	320	68	66	454; 无重复
L/H 配对分段	160	34	33	227; 无重复
双频窗口	5120	1088	1056	7264; 无重复
窗口尺寸	2×8192	2×8192	2×8192	全部一致

4.2. 整体结果与训练稳定性

四种配置在固定测试集上的三随机种子汇总结果见表 3。为减少重复展示,表中同时列出分类性能与召回均衡指标; Aug、SupCon 和 Joint 分别表示 CE + MildAug、CE + Source-aware SupCon 及联合配置。

Table 3. Main test results of four training configurations (mean $\pm$ sample standard deviation)
表 3. 四种训练配置的主要测试结果(均值 $\pm$ 样本标准差)

方法	BA	BGR	UAVR	F1	RS
CE	0.9901 ± 0.0037	0.9913 ± 0.0030	0.9888 ± 0.0053	0.9934 ± 0.0029	0.9891 ± 0.0042
Aug	0.9901 ± 0.0042	0.9878 ± 0.0108	0.9923 ± 0.0047	0.9948 ± 0.0018	0.9872 ± 0.0044
SC	0.9881 ± 0.0042	0.9878 ± 0.0060	0.9884 ± 0.0058	0.9928 ± 0.0030	0.9866 ± 0.0035
Joint	0.9926 ± 0.0051	0.9913 ± 0.0080	0.9938 ± 0.0041	0.9959 ± 0.0025	0.9910 ± 0.0050

注: μ 和 SD 分别表示均值与样本标准差, SC 表示 Source-aware SupCon。四种配置的 Accuracy 均值依次为 0.9893、0.9915、0.9883 和 0.9934, Gap 均值依次为 0.0041、0.0114、0.0060 和 0.0064。

联合配置在 Acc.、BA、UAVR、F1 和 RS 上取得最高均值, BA 较 CE 提高 0.25 个百分点, UAVR 提高 0.50 个百分点; 两者 BG R 均为 0.9913。MildAug 单独使用时减少 UAV 漏检, 但 BG R 下降并使 Gap 扩大; 单独加入 Source-aware SupCon 后, BA、RS 和 F1 均未超过 CE。联合配置的平均增益较小, 且三次运行的波动区间存在重叠, 现有结果不支持统计显著性结论。类别召回率与 Gap 见图 2。

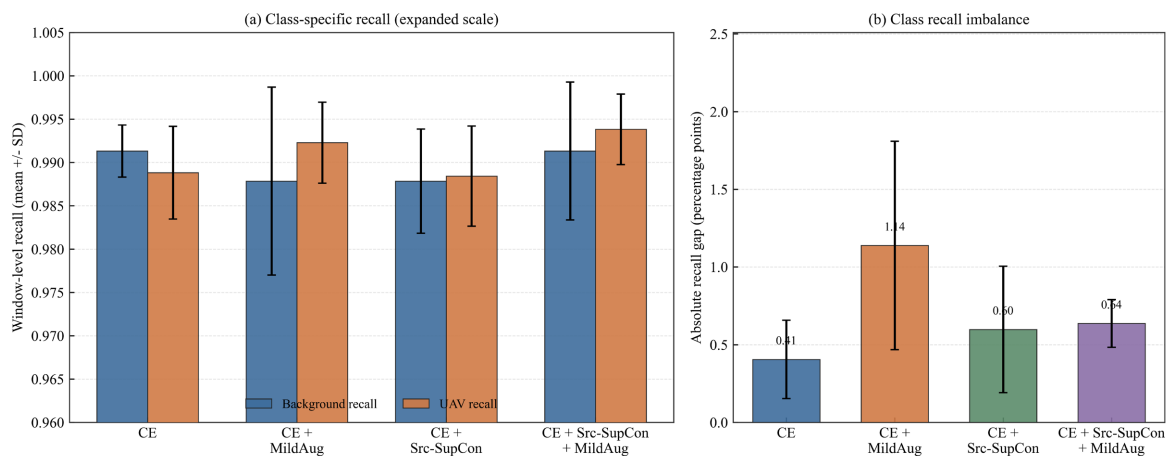


Figure 2. Recall rates and recall gap of four training configurations
图 2. 四种训练配置的分类召回率与召回率差值

三个随机种子合并统计时, CE 的窗口计数为 TN = 571、FP = 5、FN = 29、TP = 2563, 联合配置为 TN = 571、FP = 5、FN = 16、TP = 2576, 即 Background 误报数不变而 UAV 漏检减少 13 个。对每个 L/H

配对分段内 32 个窗口的概率取平均后, 四种配置均得到 $TN = 18$ 、 $FP = 0$ 、 $FN = 0$ 、 $TP = 81$ 。该汇总在当前测试清单上达到饱和, 只用于说明多窗口聚合后的判定情况, 方法差异仍以窗口结果为准, 见图 3。

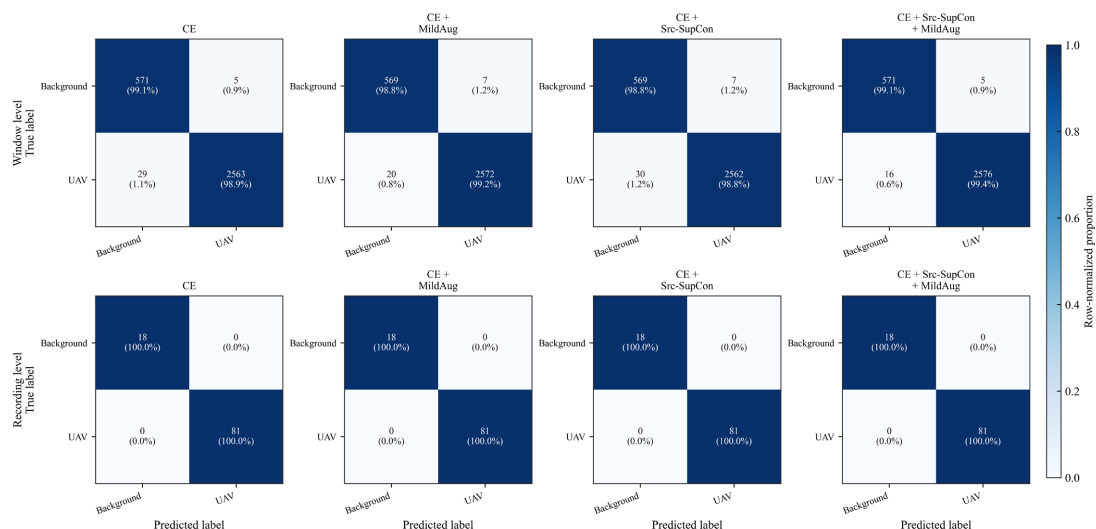


Figure 3. Pooled confusion matrices for window-based predictions and L/H paired segments

图 3. 按窗口统计及 L/H 配对分段汇总的混淆矩阵

三个随机种子下的 BA、RS、F1 与选定轮次见图 4。联合配置在种子 42、2024 和 3407 下的 BA 分别为 0.9936、0.9971 和 0.9870；结果随随机初始化变化, 因而本文同时报告均值、样本标准差和独立运行结果。选定轮次分布于第 19 至 99 轮, 统一采用验证规则选择检查点可降低收敛时点差异对比较的干扰。

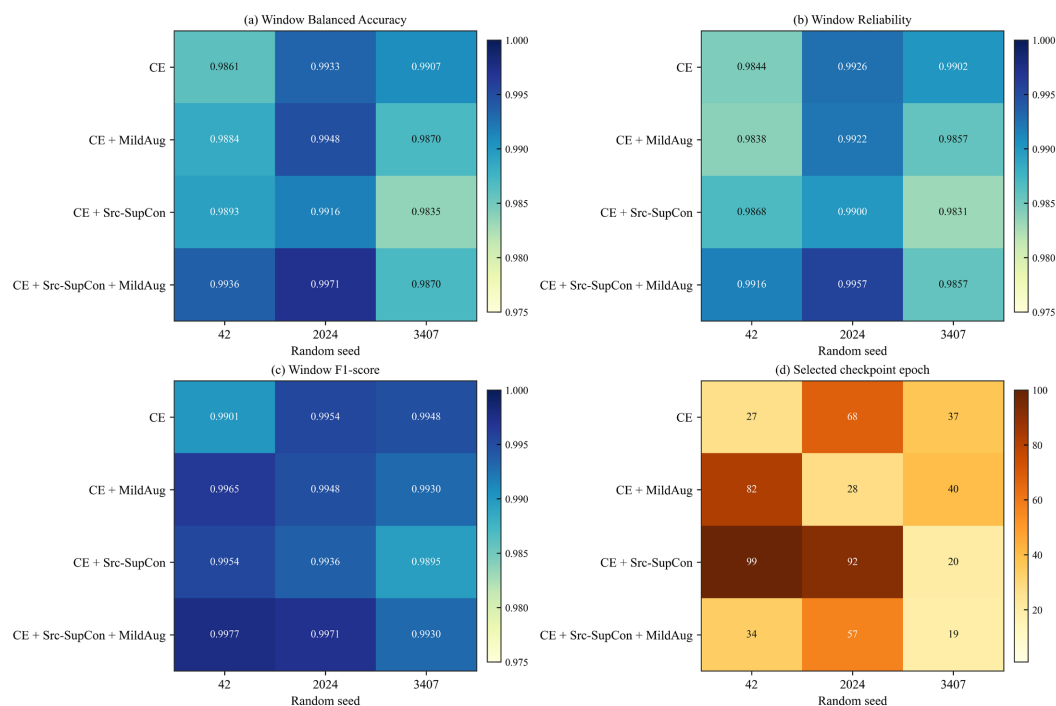


Figure 4. Evaluation results and selected epochs under different random seeds

图 4. 不同随机种子下的评价结果与模型选定轮次

图 5 给出 100 轮内训练集 BA、验证集 BA、验证集 RS 和验证交叉熵损失的三随机种子均值与样本标准差。曲线用于观察收敛与波动，最终指标均由重新载入选定检查点后在固定测试集上计算。

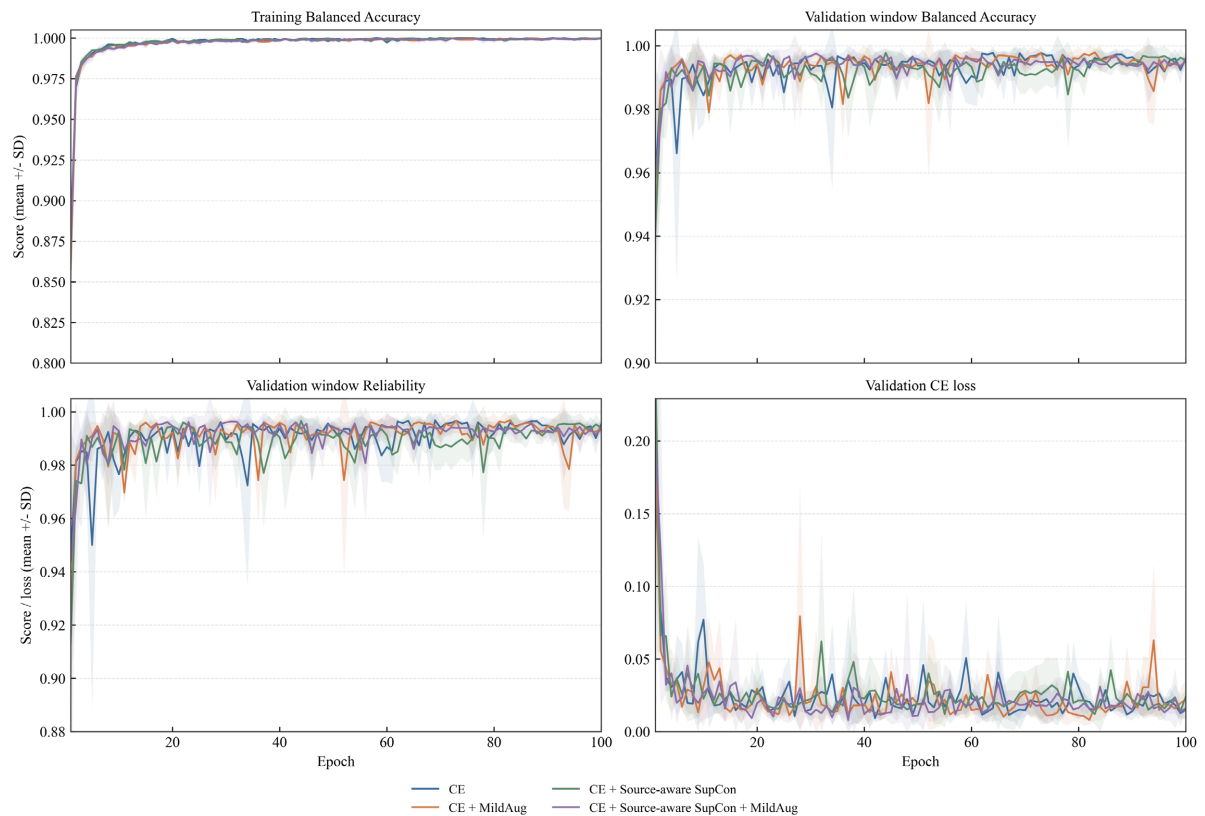


Figure 5. Training and validation curves of four training configurations

图 5. 四种训练配置的训练与验证曲线

4.3. 多表示分支消融

在固定清单、CE 训练条件和相同评价规则下，分别去除 Raw、FFT 或 STFT 分支，并对每种结构进行三随机种子训练。结果见表 4 和图 6。

Table 4. Multi-representation branch ablation results (mean $\pm$ sample standard deviation)

表 4. 多表示分支消融结果(均值 $\pm$ 样本标准差)

网络	参数量	BA	BGR	UAVR	Δ BA
完整	1,440,993	0.9901 $\pm$ 0.0037	0.9913 $\pm$ 0.0030	0.9888 $\pm$ 0.0053	0
去 Raw	988,444	0.9842 $\pm$ 0.0031	0.9826 $\pm$ 0.0030	0.9857 $\pm$ 0.0035	-0.0059
去 FFT	988,444	0.9906 $\pm$ 0.0047	0.9913 $\pm$ 0.0060	0.9900 $\pm$ 0.0059	+0.0006
去 STFT	972,684	0.9878 $\pm$ 0.0024	0.9792 $\pm$ 0.0052	0.9965 $\pm$ 0.0012	-0.0022

去除 Raw 分支后 BA 下降 0.0059，为三组消融中的最大变化；去除 STFT 分支后 BA 下降 0.0022，并表现为 UAV 召回率提高而 Background 召回率下降。去除 FFT 分支后的 BA 变化为+0.0006，小于对应样本标准差，可视为实验波动范围内的变化。由于分支去除同时改变网络参数量，上述结果仅说明当前结构中的相对变化，仍需参数量匹配实验进一步确认各表示的独立贡献。

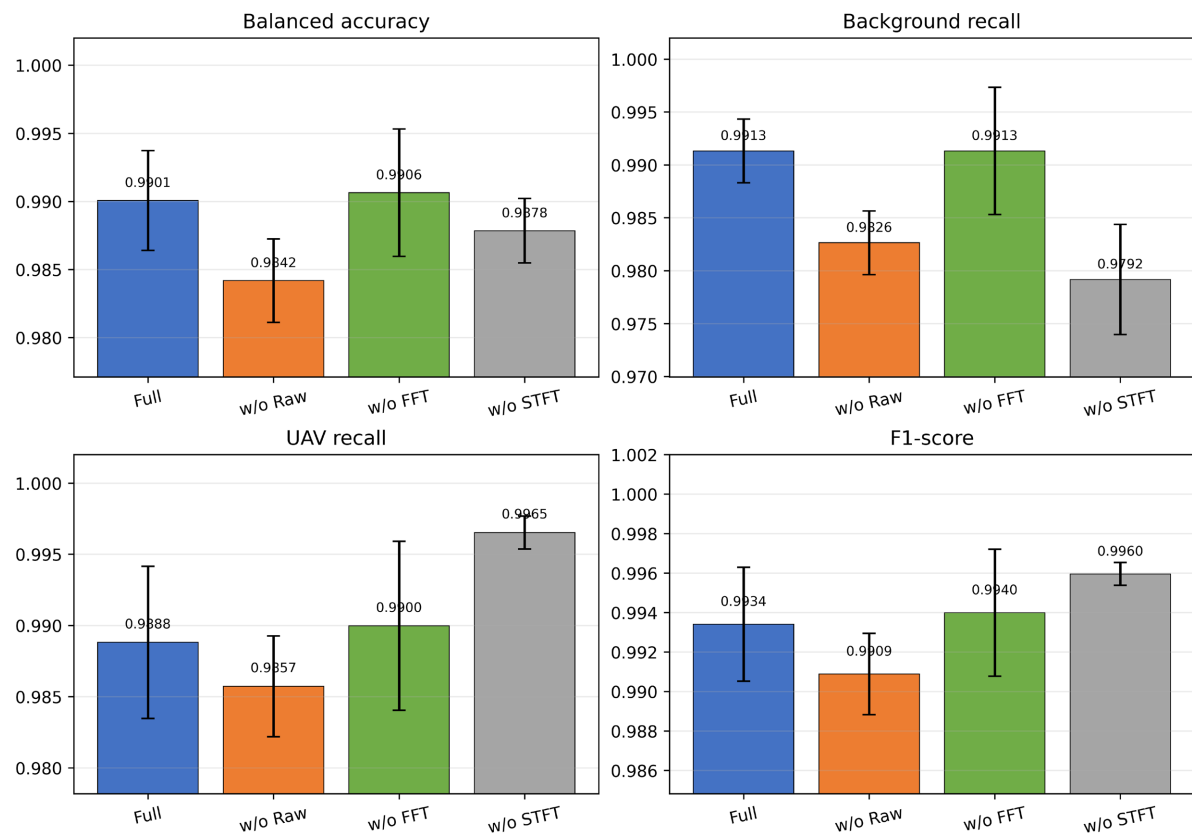


Figure 6. Leave-one-branch-out results for Raw, FFT, and STFT
图 6. Raw、FFT 与 STFT 分支消融结果

4.4. Source-Aware SupCon 参数敏感性

参数实验采用完整三支网络、固定随机种子 3407 且不启用 MildAug。 $\lambda_{ps} = 0$ 表示仅使用 CE；其余配置分别改变监督对比损失权重或温度，结果见表 5。图 7 横向子图展示两组趋势。

$\tau = 0.10$ 时， λ_{ps} 取 0.01、0.03 和 0.10 所得 BA 均低于 CE 的 0.9907，且变化不呈单调关系； $\lambda_{ps} = 0.03$ 时， $\tau = 0.20$ 的 BA 为 0.9896，最接近 CE 但仍低 0.0012。本组实验只有一个随机种子，只用于描述参数变化趋势，不能据此判断稳定性或统计显著性。

Table 5. Sensitivity to λ_{ps} and τ at random seed 3407

表 5. 随机种子 3407 下 λ_{ps} 与 τ 的参数敏感性

λ_{ps}	τ	BA	BG R	UAV R	F1
0	0.10	0.9907	0.9896	0.9919	0.9948
0.01	0.10	0.9861	0.9896	0.9826	0.9901
0.03	0.10	0.9835	0.9844	0.9826	0.9895
0.10	0.10	0.9858	0.9844	0.9873	0.9919
0.03	0.05	0.9858	0.9844	0.9873	0.9919
0.03	0.20	0.9896	0.9896	0.9896	0.9936

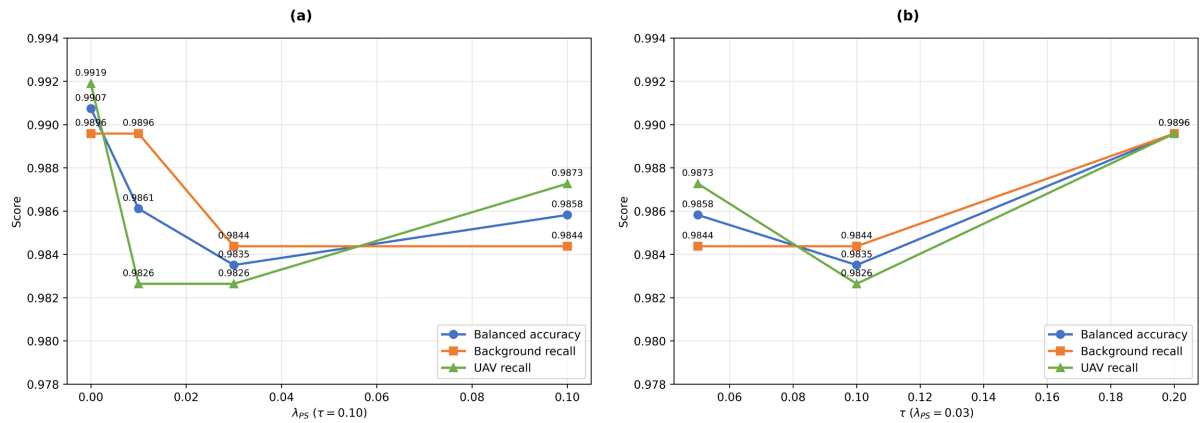


Figure 7. Sensitivity of Source-aware SupCon to (a) Loss Weight λ_{PS} and (b) Temperature τ

图 7. Source-aware SupCon 参数敏感性: (a) 损失权重 λ_{PS} ; (b) 温度 τ

4.5. 特征分布与结果讨论

为观察 Source-aware SupCon 对表示空间的影响, 从随机种子 3407 对应的 CE 与 CE + Source-aware SupCon 模型中提取固定测试集 1056 个窗口的 256 维归一化特征, 并采用相同的 `random_state = 3407`、`perplexity = 30` 和 PCA 初始化进行 t-SNE 降维, 结果见图 8。两幅子图只用于比较各自空间内的聚集与分离形态, 不对坐标位置作横向解释。

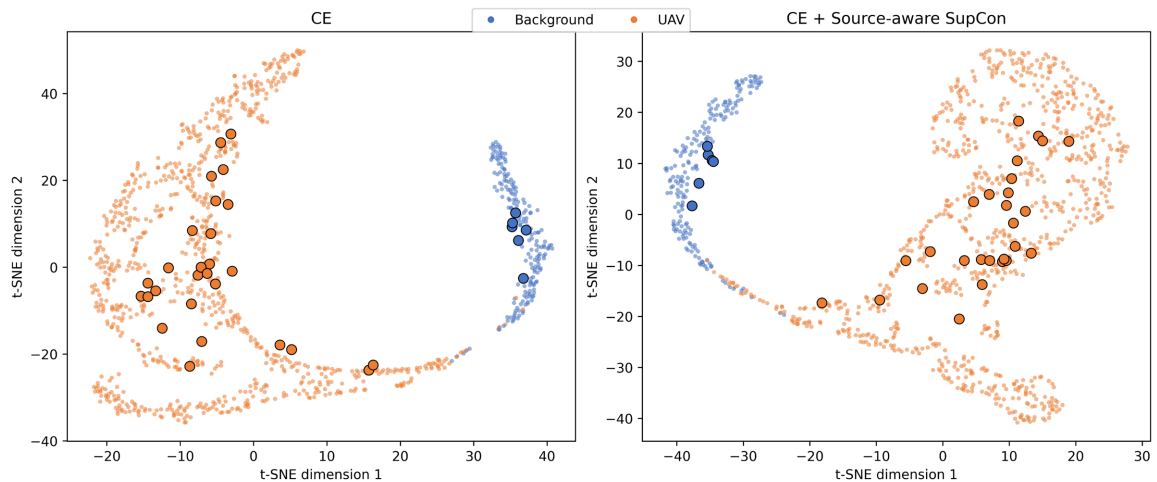


Figure 8. t-SNE distributions of test features from CE and CE plus source-aware SupCon

图 8. CE 与 CE + Source-aware SupCon 测试特征的 t-SNE 分布

引入 Source-aware SupCon 后, 同类别、不同 L/H 配对分段的平均余弦相似度由 0.9340 升至 0.9594, 不同类别相似度由 -0.9293 变为 -0.5234; 相似度间隔由 1.8632 降至 1.4828, 轮廓系数由 0.9574 降至 0.9523。该约束增强了跨配对分段的同类聚集, 同时伴随类别间隔收缩, 与其分类性能未稳定提高的结果一致。

MildAug 主要改变两类错误的分配, Source-aware SupCon 单独使用未形成稳定增益; 两者联合虽取得最高平均 BA, 但未在三个随机种子下保持一致。因此, 两项训练策略属于固定协议下的补充探索, 评价协议的有效性主要来自窗口生成前划分、固定清单和配对分段互斥。现有三随机种子结果不足以支持显著性检验, 参数敏感性结果的解释范围更为有限。

DANN 依赖明确的域标签，并通过域分类器与梯度反转层抑制域相关信息[12]；Group DRO 依赖预定义群组，并优化最不利群组损失[13]。本文的 L/H 配对分段标识仅来自文件组织关系，不能等同于设备或环境域，因此相关比较限于方法原理与适用边界。

5. 总结与展望

5.1. 研究工作总结

本文依据 BUI 活动代码和分段编号，将 454 个单频段 CSV 文件整理为 227 个 L/H 配对分段，并在窗口生成前完成 160/34/33 的训练、验证和测试划分。同一配对分段产生的窗口始终归入同一数据子集；固定数据清单、测试清单和指标规则后，对 CSV 路径、频段文件、配对分段及窗口标识进行互斥性核对。每个配对分段生成 32 个尺寸为 2×8192 的双频索引对齐窗口，共得到 7264 个窗口，由此形成可复核的实验基础。

在统一协议下，本文采用包含 1,440,993 个可训练参数的 Raw-FFT-STFT 三支网络，比较 CE、CE + MildAug、CE+Source-aware SupCon 及联合配置。三个随机种子下，联合配置的平均 BA 为 0.9926 ± 0.0051 ，较 CE 提高 0.25 个百分点，但不同随机种子的结果并不一致；四种配置对 33 个测试配对分段进行概率汇总后均正确分类，说明该汇总指标在当前测试清单上已达到饱和。

分支消融显示，去除 Raw 和 STFT 分支后 BA 分别下降 0.0059 和 0.0022；去除 FFT 分支后的变化为 +0.0006，处于实验波动范围内。Source-aware SupCon 提高了跨配对分段的同类特征相似度，同时伴随类别间隔收缩；固定随机种子下，非零对比损失权重均未超过 CE 结果。因此，本文将其视为固定协议下的探索性训练设置，不作稳定提升识别性能的结论。总体而言，本文的主要工作在于建立配对分段互斥划分、固定清单和统一评价规则，并据此限定实验结果的解释范围。

5.2. 研究局限与后续方向

DroneRF 未提供设备实例、采集日期、实验轮次及独立采集过程等信息，现有协议只能控制由文件组织关系识别的相关窗口交叉，无法验证跨设备或跨环境泛化。测试集仅含 33 个 L/H 配对分段，其中同一配对分段内的窗口仍具有相关性；主要比较使用三个随机种子，参数敏感性实验使用一个随机种子，尚不足以开展可靠的显著性检验。本文还仅考察 Background/UAV 二分类，结论不宜扩展至型号、状态或未知类别识别。

后续研究应补充设备、时间、环境和信噪比等采集信息，按独立设备或独立采集过程划分数据，并增加跨环境、跨设备及外部数据验证。模型方面，可在参数量相近条件下补充单分支与双分支组合，进一步分析 FFT 和 STFT 表示的独立贡献；针对 Source-aware SupCon，可比较不同分母掩码与正样本构造，并在多个随机种子下报告置信区间。应用研究可扩展至无人机型号与运行状态、低信噪比、频段缺失和开放集条件，同时结合参数量、推理延迟与存储开销评价部署可行性。

参考文献

- [1] Shi, X., Yang, C., Xie, W., Liang, C., Shi, Z. and Chen, J. (2018) Anti-Drone System with Multiple Surveillance Technologies: Architecture, Implementation, and Challenges. *IEEE Communications Magazine*, **56**, 68-74. <https://doi.org/10.1109/mcom.2018.1700430>
- [2] O'Shea, T.J. and Hoydis, J. (2017) An Introduction to Deep Learning for the Physical Layer. *IEEE Transactions on Cognitive Communications and Networking*, **3**, 563-575. <https://doi.org/10.1109/tccn.2017.2758370>
- [3] Al-Sa'd, M.F., Al-Ali, A., Mohamed, A., Khattab, T. and Erbad, A. (2019) RF-Based Drone Detection and Identification Using Deep Learning Approaches: An Initiative towards a Large Open Source Drone Database. *Future Generation Computer Systems*, **100**, 86-97. <https://doi.org/10.1016/j.future.2019.05.007>

-
- [4] Allahham, M.S., Al-Sa'd, M.F., Al-Ali, A., Mohamed, A., Khattab, T. and Erbad, A. (2019) DroneRF Dataset: A Dataset of Drones for RF-Based Detection, Classification and Identification. *Data in Brief*, **26**, Article ID: 104313. <https://doi.org/10.1016/j.dib.2019.104313>
 - [5] Nemer, I., Sheltami, T., Ahmad, I., Yasar, A.U. and Abdeen, M.A.R. (2021) RF-Based UAV Detection and Identification Using Hierarchical Learning Approach. *Sensors*, **21**, Article No. 1947. <https://doi.org/10.3390/s21061947>
 - [6] Shulman, D. (2026) How Much Do RF Drone Benchmarks Overstate? A Controlled Study and Theory of Data Leakage in UAV Signal Identification. <https://doi.org/10.48550/arXiv.2607.01025>
 - [7] Kapoor, S. and Narayanan, A. (2023) Leakage and the Reproducibility Crisis in Machine-Learning-Based Science. *Patterns*, **4**, Article ID: 100804. <https://doi.org/10.1016/j.patter.2023.100804>
 - [8] Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C. and Krishnan, D. (2020) Supervised Contrastive Learning. *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020*, 6-12 December 2020, 18661-18673.
 - [9] O'Shea, T.J., Corgan, J. and Clancy, T.C. (2016) Convolutional Radio Modulation Recognition Networks. In: Jayne, C. and Iliadis, L., Eds., *Engineering Applications of Neural Networks*, Springer International Publishing, 213-226. https://doi.org/10.1007/978-3-319-44188-7_16
 - [10] Mo, Y., Huang, J. and Qian, G. (2022) Deep Learning Approach to UAV Detection and Classification by Using Compressively Sensed RF Signal. *Sensors*, **22**, Article No. 3072. <https://doi.org/10.3390/s22083072>
 - [11] Cohen, L. (1995) Time-Frequency Analysis. Prentice Hall PTR.
 - [12] Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M. and Lempitsky, V. (2016) Domain-Adversarial Training of Neural Networks. *Journal of Machine Learning Research*, **17**, 1-35.
 - [13] Sagawa, S., Koh, P.W., Hashimoto, T.B. and Liang, P. (2020) Distributionally Robust Neural Networks for Group Shifts: On the Importance of Regularization for Worst-Case Generalization. *8th International Conference on Learning Representations, ICLR 2020*, Addis Ababa, 26-30 April 2020, 1-19. <https://openreview.net/forum?id=ryxGuJrFvS>