

基于多尺度时频特征与通道注意力的无人机射频指纹识别方法研究

何天洋, 郝东来, 林成琨, 索存宁

西京学院计算机学院, 陕西 西安

收稿日期: 2026年8月19日; 录用日期: 2026年9月15日; 发布日期: 2026年9月24日

摘要

针对无人机射频信号识别过程中单尺度时频表示难以兼顾不同时间-频率分辨率, 以及噪声扰动下识别性能易下降的问题, 本文提出一种基于多尺度时频特征与通道注意力的无人机射频指纹识别方法。首先, 对DroneRF公开数据集中的AR drone、Bepop drone和Phantom drone三类射频信号进行格式转换、文件级划分与固定窗口切片, 并采用不同窗口长度的短时傅里叶变换(Short-Time Fourier Transform, STFT)构建多尺度时频表示。随后, 将窗口长度为128、256和512获得的时频特征统一尺寸并进行通道拼接, 以ResNet18为主干网络, 引入压缩激励(Squeeze-and-Excitation, SE)通道注意力机制, 并结合幅值缩放、时域平移和噪声扰动等数据增强策略, 构建Multi-STFT ResNet18-SE-Aug识别模型。实验结果表明, 本文模型在文件级验证集上的识别准确率达到0.7946, Macro F1达到0.7833, 其中准确率较对应的单尺度模型提高1.65个百分点。进一步的特征域噪声扰动实验表明, 多尺度模型在不同扰动强度下均表现出更好的识别稳定性, 在10 dB和5 dB条件下准确率分别较单尺度模型提高8.42和9.81个百分点。结果表明, 多尺度时频表示能够有效融合不同时间-频率分辨率下的互补特征, 并提升无人机射频信号分类及中低信噪比特征扰动条件下的识别性能。

关键词

无人机识别, 射频指纹, 多尺度时频特征, 短时傅里叶变换, 通道注意力, 深度学习

Research on UAV Radio-Frequency Fingerprint Recognition Method Based on Multi-Scale Time-Frequency Features and Channel Attention

Tianyang He, Donglai Hao, Chengkun Lin, Cuning Suo

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

文章引用: 何天洋, 郝东来, 林成琨, 索存宁. 基于多尺度时频特征与通道注意力的无人机射频指纹识别方法研究[J]. 人工智能与机器人研究, 2026, 15(5): 1306-1319. DOI: 10.12677/airr.2026.155119

Abstract

To address the limitations of single-scale time-frequency representations in capturing complementary temporal and spectral characteristics of UAV radio-frequency (RF) signals, as well as the degradation of recognition performance under noise perturbations, this paper proposes a UAV RF fingerprint recognition method based on multi-scale time-frequency features and channel attention. First, RF signals from three UAV categories, namely AR drone, Bepop drone, and Phantom drone, are selected from the public DroneRF dataset and processed through format conversion, source-file-level partitioning, and fixed-length segmentation. Short-time Fourier transforms with window lengths of 128, 256, and 512 are then applied to construct multi-scale time-frequency representations. The resulting spectrograms are resized to a uniform resolution and concatenated along the channel dimension. ResNet18 is employed as the backbone network, while squeeze-and-excitation channel attention and data augmentation strategies, including amplitude scaling, temporal shifting, and noise perturbation, are incorporated to construct the Multi-STFT ResNet18-SE-Aug model. Experimental results show that the proposed model achieves an accuracy of 0.7946 and a Macro F1 score of 0.7833 on the file-level validation set, with an accuracy improvement of 1.65 percentage points over the corresponding single-scale model. Further feature-domain noise perturbation experiments demonstrate that the multi-scale model maintains better recognition stability under different perturbation levels, achieving accuracy improvements of 8.42 and 9.81 percentage points at 10 dB and 5 dB, respectively. These results indicate that multi-scale time-frequency representations can effectively integrate complementary RF characteristics at different time-frequency resolutions and improve UAV RF signal classification performance, particularly under medium- and low-SNR feature perturbations.

Keywords

UAV Recognition, Radio-Frequency Fingerprint, Multi-Scale Time-Frequency Features, Short-Time Fourier Transform, Channel Attention, Deep 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

1.1. 研究背景与意义

随着无人机在测绘、农业、物流和应急救援等领域的广泛应用，未经授权飞行和敏感区域入侵等问题也对低空安全提出了新的要求。现有无人机探测主要采用视觉、雷达、声学 and 射频等手段，其中射频探测能够被动获取无人机通信链路信号，不依赖目标外观和光照条件，具有较好的非合作识别潜力。

无线发射设备受功率放大器非线性、振荡器频率偏移及调制失真等硬件差异影响，其发射信号在幅度、频率和时频分布等方面表现出一定设备相关特征。深度学习能够从射频率序列或时频表示中自动学习判别信息，但单尺度时频表示难以同时兼顾局部瞬态变化和频谱结构。为此，本文研究多尺度时频特征与通道注意力结合的无人机射频识别方法，并进一步分析其在特征噪声扰动条件下的识别稳定性。

1.2. 国内外研究现状

早期无人机射频识别主要依赖人工特征和传统机器学习方法。Ezuma 等人通过能量瞬态和统计特征实现无人机射频分类[1]。随着深度学习的发展, Al-Sa'd 等人开展了基于深度网络的无人机射频检测与识别, 并形成公开 DroneRF 数据集[2] [3]; Basak 等人进一步采用深度残差网络提取无人机射频特征[4]。残差学习和 SE 通道重标定分别为深层特征提取和通道选择提供了有效结构基础[5] [6]。

针对单一信号表示信息有限的问题, Zeng 等人采用注意力机制融合不同射频特征表示[7], Zahid 等人从多尺度角度提取射频指纹信息, 并验证了其在低 SNR 条件下的有效性[8]。此外, 已有研究表明 Wi-Fi、Bluetooth 等无线干扰和 SNR 下降会明显影响无人机射频识别性能[9]。这些研究表明, 多尺度特征和注意力机制具有改善复杂射频特征表达的潜力, 但固定尺度表示、噪声扰动稳定性以及数据划分中的同源相关性仍是需要关注的问题。

1.3. 本文研究内容

本文以 DroneRF 数据集中的 AR drone、Bepop drone 和 Phantom drone 三类射频信号为研究对象, 在原始文件层面划分训练与验证数据, 并采用长度 8192 的固定窗口构建样本。在此基础上, 分别利用 128、256 和 512 三种窗口生成 STFT 时频表示, 经尺寸统一后形成三通道多尺度输入; 随后以 ResNet18 为主干网络, 引入 SE 通道注意力和数据增强策略, 构建 Multi-STFT ResNet18-SE-Aug 模型。

实验从网络结构、单尺度与多尺度特征、分类混淆及特征域噪声扰动等方面进行评价。最终模型在文件级验证集上取得 0.7946 的 Accuracy 和 0.7833 的 Macro F1, 并在中低 SNR 特征扰动条件下表现出更好的识别稳定性。此外, 本文通过留一类实验对未知无人机拒识能力进行了初步探索。

2. 数据集与射频信号特征构建

本章介绍本文采用的 DroneRF 数据集及射频信号预处理方法, 包括文件级划分、固定窗口切片、STFT 时频表示、多尺度时频特征构建和数据增强, 为后续识别模型提供统一的数据与特征输入。

2.1. DroneRF 数据集介绍

本文采用公开 DroneRF 数据集开展实验[3], 选取 AR drone、Bepop drone 和 Phantom drone 三类无人机的射频信号作为识别对象。为减少同一原始记录中相邻切片同时进入训练集和验证集造成的数据相关性泄漏, 本文首先在原始文件层面完成数据划分, 再分别生成训练和验证切片。最终构建 12,600 个样本, 其中训练集 10,000 个、验证集 2600 个, 具体分布见表 1。

Table 1. Dataset class distribution and sample partition
表 1. 数据集类别分布与样本划分

类别	训练集样本数	验证集样本数	类别说明
AR drone	3200	1000	无人机类别 1
Bepop drone	3400	800	无人机类别 2
Phantom drone	3400	800	无人机类别 3
合计	10,000	2600	共三类无人机

2.2. 射频信号数据预处理

DroneRF 原始射频序列数据量较大, 为提高训练阶段的随机读取效率, 本文将 CSV 文件转换为 float32

格式的 NPY 文件，并建立包含文件路径、切片位置和类别标签的索引。各切片在进行时频变换前采用均值 - 方差标准化，以降低不同样本幅值分布差异的影响。

数据预处理主要包括三个步骤。首先，对原始射频文件进行类别整理，根据文件所属无人机类别建立标签映射关系，将 AR drone、Bepop drone 和 Phantom drone 分别映射为 0、1、2 三个类别标签。其次，将原始长序列转换为统一的数据格式，并保存为 NPY 文件，提高后续切片读取和批量加载效率。最后，构建样本索引文件，记录每个切片样本对应的文件路径、起始位置、结束位置和类别标签，为模型训练阶段的数据加载提供统一入口。

对于每一个切片样本，本文在构建时频图之前进行标准化处理，以降低不同文件或不同片段之间幅值分布差异对模型训练的影响。设原始切片信号为 $x = \{x_1, x_2, \dots, x_L\}$ ，其中 L 表示切片长度，本文采用如下标准化方式：

$$\hat{x}_i = \frac{x_i - \mu}{\sigma + \varepsilon} \quad (1)$$

其中， μ 和 σ 分别表示当前切片信号的均值和标准差， ε 为防止分母为零而设置的极小常数。通过该处理，不同样本的幅值分布被映射到相对一致的范围，有助于模型更加关注信号的结构性特征，而不是绝对幅值差异。

2.3. 固定窗口切片与样本构建

原始射频信号通常是连续采集得到的长时间序列，而深度学习模型需要固定尺寸的输入。因此，本文采用固定窗口切片方法，将长序列划分为长度一致的短时间段。设原始射频频序列为 x ，切片长度为 L ，则第 i 个切片样本可表示为：

$$x_i = x[iL : (i+1)L] \quad (2)$$

本文将切片长度设置为 8192，并对每个原始文件最多抽取 200 个切片，以兼顾信号信息保留与样本规模控制。数据划分在切片生成前完成，共 63 个原始文件，其中 50 个用于训练、13 个用于验证，最终得到 10,000 个训练样本和 2600 个验证样本。由于训练集与验证集来源于不同原始文件，该策略能够减少相邻切片跨数据子集造成的数据相关性泄漏。原始射频文件经过格式转换、文件级划分和固定窗口切片后，生成用于模型训练和验证的标准化样本索引。

2.4. 单尺度 STFT 时频表示

为同时刻画无人机射频信号的时间变化和频率结构，本文采用短时傅里叶变换(STFT)将一维射频频片段转换为二维时频表示[10]。STFT 通过滑动窗口对局部信号进行傅里叶分析，其表达式为：

$$X(t, f) = \sum_{n=-\infty}^{+\infty} x[n] w[n-t] e^{-j2\pi f n} \quad (3)$$

其中， $x[n]$ 表示输入射频信号， $w[n-t]$ 表示以时间位置 t 为中心的窗函数， f 表示频率变量， $X(t, f)$ 表示信号在时间和频率维度上的局部频谱分布。STFT 能够同时保留信号的时间变化信息和频率结构信息，适合用于无人机射频信号的特征表达。

由于 STFT 结果为复数形式，本文取其幅值并进行对数变换，得到用于模型输入的时频幅值图：

$$(t, f) = \log(1 + |X(t, f)|) \quad (4)$$

其中， $S(t, f)$ 表示对数幅值时频图。对数变换能够压缩幅值动态范围，减弱极大值对模型训练的影响，使

时频图中的弱特征结构更加容易被模型学习。单尺度 STFT 最终得到的输入尺寸为 $1 \times 129 \times 65$ ，其中第一个数值 1 表示通道数(单通道输入)，129 表示频率维度的大小，65 表示时间维度的大小。

图 1 展示了对应的 STFT 时频图示例。从图中可以看出，经过 STFT 变换后，原始一维序列被转换为二维时频结构，信号的局部频率变化和时频纹理特征更加直观。

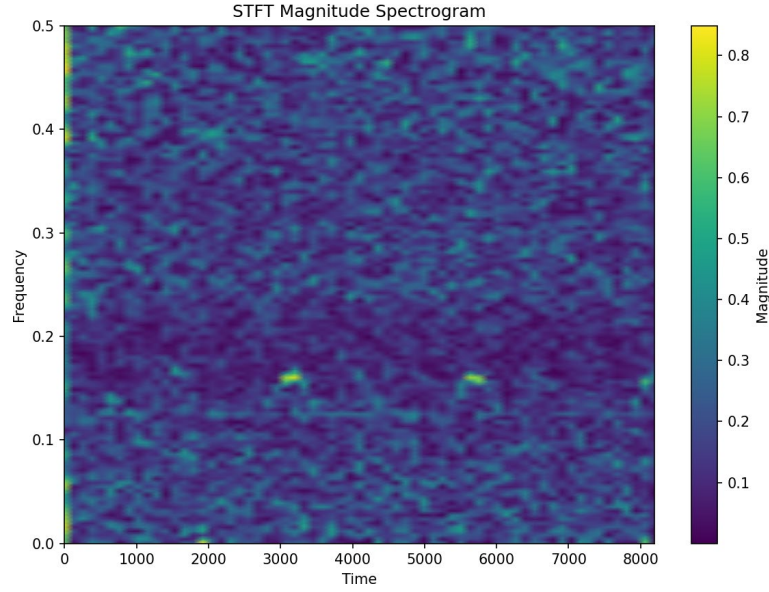


Figure 1. Example of a single-scale STFT spectrogram
图 1. 单尺度 STFT 时频图示例

2.5. 多尺度 STFT 时频特征构建

单尺度 STFT 虽然能够将射频信号转换为时频图,但其时间分辨率和频率分辨率受到窗口长度影响。一般而言,较短的窗口能够更好地捕捉信号的局部瞬态变化,但频率分辨率相对较低;较长的窗口能够提供更精细的频率结构,但对时间变化的刻画能力较弱。因此,单一窗口长度难以同时兼顾不同尺度下的射频特征表达。

为增强模型对复杂射频信号的建模能力,本文在单尺度 STFT 基础上进一步构建多尺度 STFT 特征。具体而言,对于同一段长度为 8192 的射频切片,分别采用不同窗口长度进行 STFT 计算,得到多个尺度下的时频表示。本文选取窗口长度为 128、256 和 512 的三种 STFT 尺度,分别记为(S_{128})、(S_{256})和(S_{512})。随后将不同尺度得到的时频图统一调整为相同尺寸,并在通道维度进行拼接,形成三通道多尺度时频特征:

$$S_{\text{multi}} = \text{Concat}(S_{128}, S_{256}, S_{512}) \quad (5)$$

其中, S_{multi} 表示多尺度 STFT 特征, $\text{Concat}()$ 表示通道维度拼接操作。经过该处理后,模型输入由原来的 $1 \times 129 \times 65$ 扩展为 $3 \times 129 \times 65$ 。其中,三个通道分别对应短窗口、中窗口和长窗口下的时频表示。

为保证不同尺度时频特征构建过程的一致性,本文对三种 STFT 尺度采用统一的参数设置原则。设 STFT 窗口长度为 N_w ,则窗口重叠长度设置为窗口长度的 50%,相邻窗口之间的移动步长为 $N_w/2$,FFT 点数设置为与当前窗口长度一致。

对于 STFT 得到的复数时频系数 $Z(t, f)$,本文首先计算其幅值,并采用对数变换压缩不同频谱能量之间的动态范围,其表达式为:

$$S(t, f) = \log(1 + |Z(t, f)|) \quad (6)$$

其中, $S(t, f)$ 表示经过对数压缩后的时频幅值特征。

随后, 对每个尺度得到的时频特征分别进行均值 - 方差标准化, 并采用双线性插值将其统一调整为 129×65 的二维尺寸。完成上述处理后, 将三个尺度的时频特征沿通道维度进行拼接, 最终形成尺寸为 $3 \times 129 \times 65$ 的多尺度时频输入。

三种尺度分别提供局部瞬态、稳定时频纹理和较细频率结构等互补信息, 为后续深度网络的联合特征学习提供基础。通过三种尺度的联合表达, 模型能够获得比单尺度输入更丰富的射频指纹信息, 从而提升无人机类别识别能力和复杂噪声环境下的鲁棒性。

图 2 展示了本文多尺度 STFT 特征构建流程。对于同一射频切片, 本文分别计算三种不同窗口长度下的 STFT 时频图, 并将其统一尺寸后拼接为三通道输入, 为后续 ResNet18-SE 识别模型提供多尺度时频特征。

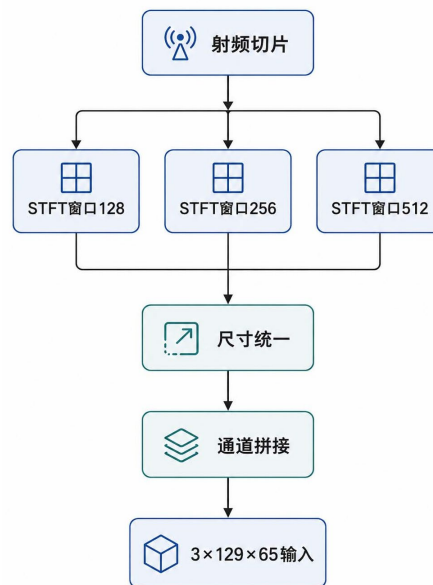


Figure 2. Multi-scale STFT feature construction workflow

图 2. 多尺度 STFT 特征构建流程图

表 2 给出了本文主要数据预处理与时频特征构建参数

Table 2. Parameters for data preprocessing and time-frequency feature construction

表 2. 数据预处理与时频特征构建参数

参数名称	参数取值	说明
切片长度	8192	每个样本包含的时域采样点数
每文件最大切片数	200	控制样本规模并保持文件贡献均衡
训练集样本数	10,000	文件级划分后生成
验证集样本数	2600	文件级划分后生成
单尺度输入尺寸	$1 \times 129 \times 65$	单通道 STFT 时频图

续表

多尺度窗口长度	128、256、512	三种 STFT 尺度
多尺度输入尺寸	$3 \times 129 \times 65$	三通道时频图输入
特征归一化	均值方差标准化	减小样本幅值分布差异

2.6. 数据增强策略

为提高模型对信号扰动的适应能力，本文在训练阶段对射频切片进行随机时域平移和噪声扰动，以增加训练样本的变化形式。验证阶段不进行数据增强，以保证评估结果能够反映模型在未增强样本上的识别性能。设原始切片信号为 $x[n]$ ，增强后的信号可表示为：

$$x_{\text{aug}}[n] = ax[(n - \tau) \bmod L] + \eta[n] \quad (7)$$

其中， a 表示幅值缩放因子， τ 表示时域平移量， L 表示切片长度， $\eta[n]$ 表示附加噪声。

本章完成了 DroneRF 数据整理、文件级样本划分以及单尺度和多尺度 STFT 特征构建，为后续识别模型训练提供统一输入。

3. 基于多尺度时频特征与通道注意力的识别方法

本章在多尺度 STFT 特征基础上构建无人机射频识别模型，以 ResNet18 为主干网络，引入 SE 通道注意力机制，并结合多尺度时频输入完成类别预测。

3.1. 整体方法框架

本文方法包括射频信号切片、多尺度时频特征构建、深度特征提取和分类识别四个阶段。首先将原始射频序列切分为长度 8192 的片段，分别采用 128、256 和 512 三种窗口进行 STFT 变换，并将时频图统一为 129×65 后沿通道维度拼接为 $3 \times 129 \times 65$ 输入；随后由 ResNet18-SE 提取深层特征，经全连接层和 Softmax 输出无人机类别概率。

设多尺度时频输入为 S_{multi} ，深度特征提取网络为 $f_{\theta}(\cdot)$ ，则模型提取到的深层特征可表示为：

$$h = f_{\theta}(S_{\text{multi}}) \quad (8)$$

其中， θ 表示模型参数， h 表示经过深度卷积网络提取到的特征向量。分类层进一步将特征映射为类别得分：

$$o = Wh + b \quad (9)$$

其中， W 和 b 分别表示全连接层的权重和偏置， o 表示模型对各类别的原始输出得分。对于三分类任务，本文将无人机类别数设为 $C = 3$ ，分别对应 AR drone、Bepop drone 和 Phantom drone。经过 Softmax 函数后，第 c 类的预测概率为：

$$p_c = \frac{\exp(o_c)}{\sum_{k=1}^C \exp(o_k)} \quad (10)$$

最终模型选择概率最大的类别作为预测结果：

$$\hat{y} = \arg \max_c p_c \quad (11)$$

上述流程实现了从原始射频切片到无人机类别标签的端到端识别。

3.2. 基线识别模型设计

为了验证不同网络结构对无人机射频时频特征学习能力的影响，本文首先构建了浅层卷积网络

SimpleCNN 和深层残差网络 ResNet18 两类基线模型[5]。SimpleCNN 作为浅层基线，用于验证 STFT 时频特征的基本可分类性；ResNet18 则用于分析深层残差结构对时频纹理特征提取的作用。对于输入特征 x ，残差单元的输出可表示为：

$$y = \mathcal{F}(x; W) + x \quad (12)$$

其中， $\mathcal{F}(x; W)$ 表示残差分支学习到的非线性映射， W 表示该残差分支中的卷积参数， x 表示恒等映射分支。残差连接使网络不必直接学习完整映射，而是学习输入与目标输出之间的残差变化，有助于提高深层网络的训练稳定性。

针对本文 STFT 时频图尺寸较小以及多尺度三通道输入的特点，本文在标准 ResNet18 基础上对输入端结构进行适应性调整。模型不采用 ImageNet 预训练权重，而是从随机初始化参数开始训练。对于单尺度 STFT 输入，第一层卷积的输入通道数设置为 1；对于多尺度 STFT 输入，则调整为 3。与此同时，将原 ResNet18 输入端的大卷积核调整为 3×3 卷积，步长设置为 1，填充设置为 1，并取消输入端最大池化操作，以避免在网络浅层阶段对 129×65 的时频图进行过快下采样，从而尽可能保留局部频谱纹理和短时变化信息。

在上述调整基础上，网络仍保留 ResNet18 的四级残差特征提取结构。多尺度输入依次经过第一层卷积以及 Layer1、Layer2、Layer3 和 Layer4 进行层次化特征提取，并在最终全局平均池化后通过全连接层输出三类无人机的分类结果。

3.3. ResNet18-SE 注意力识别模型

为增强网络对关键通道的选择能力，在 ResNet18 各残差阶段后引入 SE 通道注意力模块[6]。

SE 模块主要包括 Squeeze、Excitation 和 Scale 三个阶段，首先利用全局平均池化获得通道描述向量，再通过两层全连接网络生成通道权重，并对原始特征进行重标定。在 Excitation 阶段，SE 模块通过两层全连接网络学习不同通道之间的依赖关系，并生成通道权重：

$$s = \sigma(W_2 \delta(W_1 z)) \quad (13)$$

其中， W_1 和 W_2 表示两层全连接层的权重参数， $\delta(\cdot)$ 表示 ReLU 激活函数， $\sigma(\cdot)$ 表示 Sigmoid 函数， s 表示学习得到的通道权重向量。该权重向量的每一个元素均位于 0 到 1 之间，用于表示对应通道的重要程度。最后，在 Scale 阶段，SE 模块利用通道权重对原始特征进行重标定：

$$\tilde{u}_c(i, j) = s_c \cdot u_c(i, j) \quad (14)$$

其中， $\tilde{u}_c(i, j)$ 表示重标定后的第 c 个通道特征， s_c 表示该通道对应的注意力权重。通过该方式，模型能够增强对识别任务更有贡献的通道响应，同时抑制冗余或干扰性特征。该模块可以灵活嵌入残差网络不同阶段，用于增强网络对关键频谱通道的建模能力。

本文分别在 ResNet18 的 Layer1、Layer2、Layer3 和 Layer4 四个残差特征提取阶段之后引入 SE 模块，对不同层级输出的特征通道进行自适应重标定，SE 模块的通道压缩比例 r 设置为 16，通过在不同残差阶段逐级进行通道特征校准，模型能够在浅层局部纹理与深层语义特征中分别强化与无人机类别判别相关的有效响应。

3.4. 多尺度时频特征融合方法

本文将窗口长度 128、256 和 512 获得的三种 STFT 时频图统一尺寸后沿通道维度拼接，形成多尺度输入，并由 ResNet18-SE 自动学习不同尺度之间的互补关系。对于同一射频切片 $x[n]$ ，本文分别计算三种窗口长度下的 STFT 时频图，经过尺寸统一后，在通道维度拼接为多尺度输入：

$$S_{\text{multi}} = \text{Concat}(S_{128}, S_{256}, S_{512}), S_{\text{multi}} \in \mathbb{R}^{3 \times 129 \times 65} \quad (15)$$

其中，三个通道分别对应不同窗口尺度下的时频表示，不同尺度对应不同的时间-频率分辨率，能够为网络提供互补的局部变化与频谱结构信息。

在模型结构上，多尺度输入并不需要改变后续分类流程，只需要将 ResNet18-SE 的第一层卷积输入通道数由 1 调整为 3。随后，卷积层会在空间维度和通道维度上共同学习多尺度特征之间的关联关系。与手工设计特征融合规则相比，该方式能够通过网络训练自动学习不同尺度特征的贡献程度。图 3 展示了本文多尺度时频特征融合与识别模型结构。

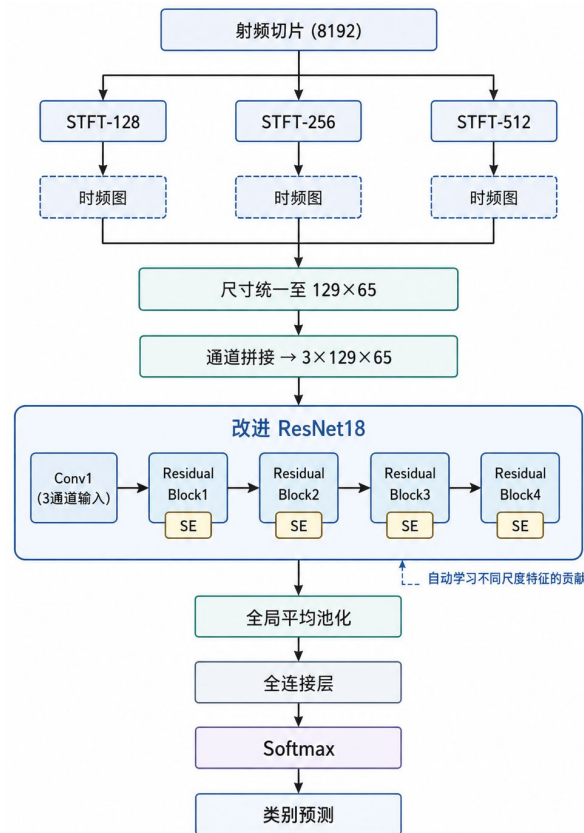


Figure 3. Multi-scale STFT and ResNet18-SE fusion recognition architecture
图 3. 多尺度 STFT 与 ResNet18-SE 融合识别结构

通过上述设计，本文最终模型能够同时利用多尺度时频表示、残差深层特征提取和通道注意力机制，从而形成完整的无人机射频指纹识别方法。

3.5. 模型训练与优化策略

本文将无人机射频指纹识别任务建模为三分类问题。对于每个输入样本，模型输出三类无人机的预测概率，并采用交叉熵损失函数进行训练。设训练集中共有 N 个样本，类别数为 C ，样本 i 的真实标签为 $y_{i,c}$ ，模型预测概率为 $\hat{y}_{i,c}$ ，则交叉熵损失函数定义为：

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c}) \quad (16)$$

其中, $y_{i,c}$ 为 one-hot 编码形式的真实标签。当样本 i 属于类别 c 时, $y_{i,c} = 1$, 否则 $y_{i,c} = 0$ 。交叉熵损失能够衡量模型预测概率分布与真实标签分布之间的差异, 是多分类任务中常用的优化目标。

模型训练采用 Adam 优化器。Adam 能够根据一阶矩估计和二阶矩估计自适应调整参数更新步长, 适合深度网络训练[11]。本文训练阶段采用数据增强策略, 包括幅值缩放、时域平移和噪声扰动。该策略已在前文, 进行介绍, 其目的是增加训练样本的扰动形式, 使模型对实际采集环境中的幅值变化、局部偏移和噪声干扰具有更强适应能力。需要注意的是, 数据增强仅用于训练集, 验证集不进行增强处理。训练过程中, 模型以验证集准确率作为最佳模型保存依据。该策略能够避免直接使用最后一轮模型带来的过拟合风险。表 3 给出了本文最终模型的主要训练参数设置。

Table 3. Training parameter settings of the proposed model

表 3. 模型训练参数设置

参数名称	参数取值	说明
输入尺寸	$3 \times 129 \times 65$	多尺度 STFT 三通道输入
类别数	3	AR drone, Bepop drone, Phantom drone
Batch size	64	每批训练样本数
Optimizer	Adam	模型参数优化器
Learning rate	$1e-4$	初始学习率
Weight decay	$1e-5$	权重衰减系数
Loss function	Cross Entropy	多分类交叉熵损失
Epoch	15	最大训练轮次
数据增强	是	幅值扰动、时域平移、噪声扰动
最佳模型选择	验证集 Accuracy 最高	用于最终测试与分析

4. 实验结果与分析

本章基于第 2 章构建的无人机射频时频特征和第 3 章提出的 Multi-STFT ResNet18-SE-Aug 识别模型, 对不同模型的性能、训练过程、类别识别差异以及噪声鲁棒性进行实验验证。实验主要围绕三类无人机射频信号展开, 包括 AR drone、Bepop drone 和 Phantom drone。通过准确率、精确率、召回率、F1-score、混淆矩阵以及不同信噪比条件下的鲁棒性测试, 综合分析本文方法的有效性。

4.1. 实验环境与评价指标

本文实验在 Windows 下完成, 使用 Python 3.10 和 PyTorch 深度学习框架搭建模型, 训练过程采用 CUDA 进行加速。实验数据集采用文件级划分方式构建, 其中训练集包含 10,000 个样本, 验证集包含 2600 个样本。模型输入为 STFT 时频图, 单尺度模型输入尺寸为 $1 \times 129 \times 65$, 多尺度模型输入尺寸为 $3 \times 129 \times 65$ 。

本文采用 Accuracy、Precision、Recall、F1-score、Macro F1 和 Weighted F1 评价模型性能。其中 Accuracy 反映总体分类正确率, Macro F1 对各类别赋予相同权重, 用于评价类别间识别均衡性。

4.2. 不同模型识别性能对比

为了验证不同网络结构和特征构建方式对无人机射频指纹识别性能的影响, 本文依次构建了 SimpleCNN、ResNet18、ResNet18-Regularized、ResNet18-SE、ResNet18-SE-Aug 和 Multi-STFT ResNet18-

SE-Aug 六组模型。其中, SimpleCNN 作为浅层卷积基线模型, ResNet18 用于验证残差深层网络的有效性, ResNet18-SE 用于验证通道注意力机制的作用, ResNet18-SE-Aug 用于分析数据增强对模型泛化能力的影响, Multi-STFT ResNet18-SE-Aug 为本文最终模型。表 4 给出了不同模型在验证集上的性能对比结果。

Table 4. Performance comparison of different recognition models
表 4. 不同识别模型性能对比

模型	输入特征	主要策略	Accuracy	Macro F1
SimpleCNN	Single-STFT	浅层卷积基线	0.5200	0.4700
ResNet18	Single-STFT	残差网络	0.7700	0.7600
ResNet18-Regularized	Single-STFT	正则化约束	0.7100	0.7000
ResNet18-SE	Single-STFT	SE 通道注意力	0.7800	0.7600
ResNet18-SE-Aug	Single-STFT	SE 注意力与数据增强	0.7781	0.7700
Multi-STFT ResNet18-SE-Aug	Multi-STFT	多尺度时频特征、SE 注意力与数据增强	0.7946	0.7833

表 4 显示, SimpleCNN 的 Accuracy 仅为 0.52, 采用 ResNet18 后提升至 0.77, 说明深层残差结构明显增强了时频特征表达能力。进一步加入 SE 注意力和数据增强后, 模型在准确率和类别均衡性方面得到改善。最终 Multi-STFT ResNet18-SE-Aug 取得最高 Accuracy 0.7946 和 Macro F1 0.7833, 表明多尺度时频表示能够进一步提升整体识别性能。

4.3. 多尺度时频特征有效性分析

为了进一步分析多尺度 STFT 特征的作用, 本文重点比较单尺度 ResNet18-SE-Aug 与 Multi-STFT ResNet18-SE-Aug 的识别结果。两种模型均采用 ResNet18-SE 作为主干网络, 并在训练阶段引入相同的数据增强策略, 区别在于前者使用单尺度 STFT 输入, 后者使用由窗口长度 128、256 和 512 构成的三通道多尺度 STFT 输入。

单尺度 STFT 只能在固定窗口长度下描述信号的时频结构, 其时间分辨率和频率分辨率相对固定, 难以同时兼顾局部瞬态变化和稳定频谱结构。多尺度 STFT 则能够从多个窗口尺度对同一射频切片进行表征, 其中短窗口更适合捕捉局部突变, 中窗口有利于保留稳定时频纹理, 长窗口更有助于刻画频率分布和谱线结构。通过通道拼接后, 模型能够在卷积过程中自动学习不同尺度特征之间的互补关系。

多尺度模型在 Accuracy、Macro Precision、Macro Recall 和 Macro F1 等指标上均优于单尺度模型, 说明不同 STFT 尺度提供的互补信息能够改善整体及分类别识别效果。

4.4. 分类报告与混淆矩阵分析

为了进一步分析最终模型对不同无人机类别的识别能力, 本文对 Multi-STFT ResNet18-SE-Aug 模型在验证集上的分类结果进行统计。图 4 展示了最终 Multi-STFT ResNet18-SE-Aug 模型的原始混淆矩阵和按真实类别归一化后的混淆矩阵。可见, Bepop drone 识别效果最好, Recall 达到 0.9475; AR drone 的 Recall 为 0.8530。相比之下, Phantom drone 的 Recall 仅为 0.5687, 其中 205 个样本被误判为 AR drone、140 个被误判为 Bepop drone。值得注意的是, Bepop drone 与 Phantom drone 在验证集中均包含 800 个样本, 但二者 Recall 差异明显, 因此 Phantom 识别困难不能简单归因于验证样本数量。结合其较高 Precision (0.8067)和较低 Recall 可以判断, 模型对 Phantom 的判定范围相对保守, 部分 Phantom 时频特征与另外两类存在较明显重叠, 导致样本被吸收到 AR 和 Bepop 决策区域。

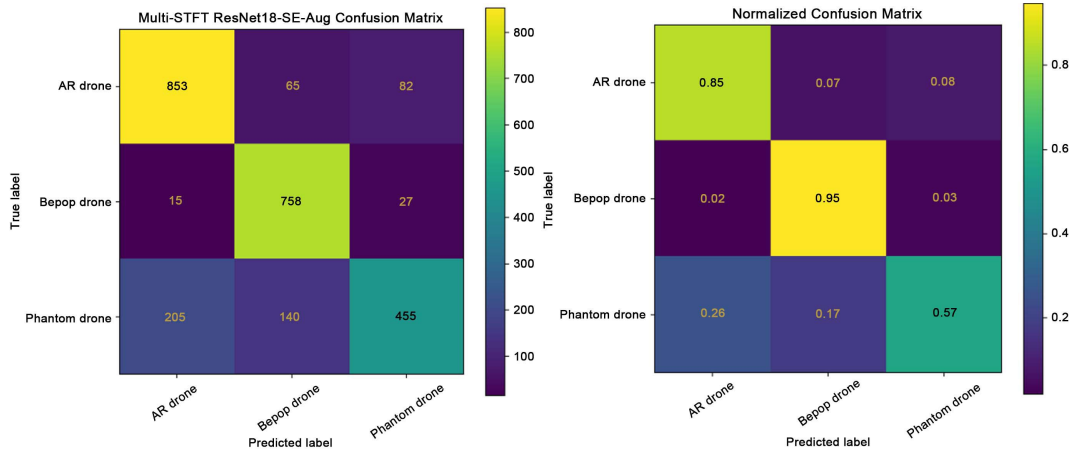


Figure 4. Confusion matrix and normalized confusion matrix of the Multi-STFT ResNet18-SE-Aug model

图 4. Multi-STFT ResNet18-SE-Aug 模型混淆矩阵与归一化混淆矩阵

后续可通过特征空间可视化、类别中心距离分析及度量学习进一步验证该现象，并增强易混淆类别的类间分离能力。

4.5. 不同信噪比条件下的鲁棒性分析

在实际低空安全应用场景中，无人机射频信号容易受到环境噪声、传输衰减和复杂电磁干扰影响。为了验证本文模型在噪声干扰条件下的识别稳定性，本文进一步开展不同信噪比条件下的鲁棒性实验。本文在标准化后的 STFT 时频特征张量上注入不同强度的高斯噪声，并根据特征功率与噪声功率之比设置 clean、20 dB、10 dB、5 dB 和 0 dB 五种条件。设原始输入特征为 x ，加入噪声后的输入为：

$$x_{\text{noise}} = x + \eta \quad (17)$$

其中， η 表示附加高斯噪声。信噪比可表示为：

$$\text{SNR} = 10 \log_{10} \frac{P_{\text{signal}}}{P_{\text{noise}}} \quad (18)$$

其中， P_{signal} 表示信号功率， P_{noise} 表示噪声功率。SNR 数值越低，表示噪声干扰越强，模型识别难度越大。

图 5 展示了不同 SNR 条件下单尺度模型与多尺度模型的鲁棒性对比曲线。

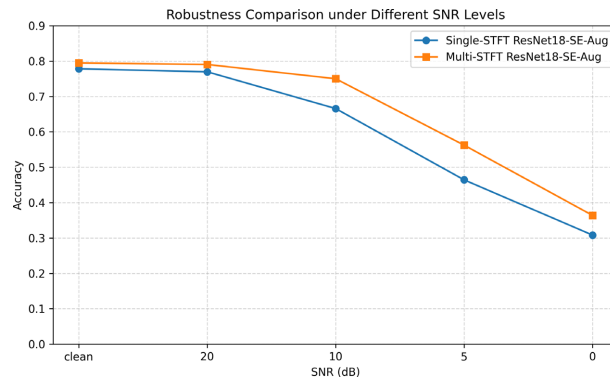


Figure 5. Robustness comparison of single-scale and multi-scale models under different SNR levels

图 5. 不同 SNR 条件下单尺度与多尺度模型鲁棒性对比

随着 SNR 降低, 两种模型的识别准确率均明显下降, 但 Multi-STFT 模型在所有扰动条件下均优于单尺度模型。其中, 在 10 dB 和 5 dB 条件下分别提升 0.0842 和 0.0981, 优势最为明显, 说明多尺度特征在中低信噪比特征扰动条件下具有较强的互补作用。0 dB 条件下两种模型性能均明显下降, 表明强扰动仍会显著影响识别效果。综上, 多尺度 STFT 方法不仅提升了无人机射频指纹识别的整体性能, 也增强了模型对噪声干扰的适应能力, 是本文方法的重要优势之一。

4.6. 开集识别初步实验分析

此外, 本文采用留一类方式初步分析未知目标拒识能力, 以 AR drone 和 Bepop drone 作为已知类、Phantom drone 作为未知类, 并根据样本特征到已知类别原型的距离进行阈值判定。设样本特征为 h , 第 c 个已知类别的特征原型为 μ_c , 则样本到最近已知类别原型的距离可表示为:

$$d(h) = \min_c \|h - \mu_c\|_2 \quad (19)$$

当距离超过阈值 T 时, 样本被判定为未知类:

$$\hat{y} = \begin{cases} \arg \min_c \|h - \mu_c\|_2, & d(h) \leq T \\ \text{unknown}, & d(h) > T \end{cases} \quad (20)$$

实验中已知类接受样本的分类 Accuracy 达到 0.9564, 但未知类拒识率仅为 0.1225。已知类与未知类样本到最近类别原型的距离分布如图 6 所示。

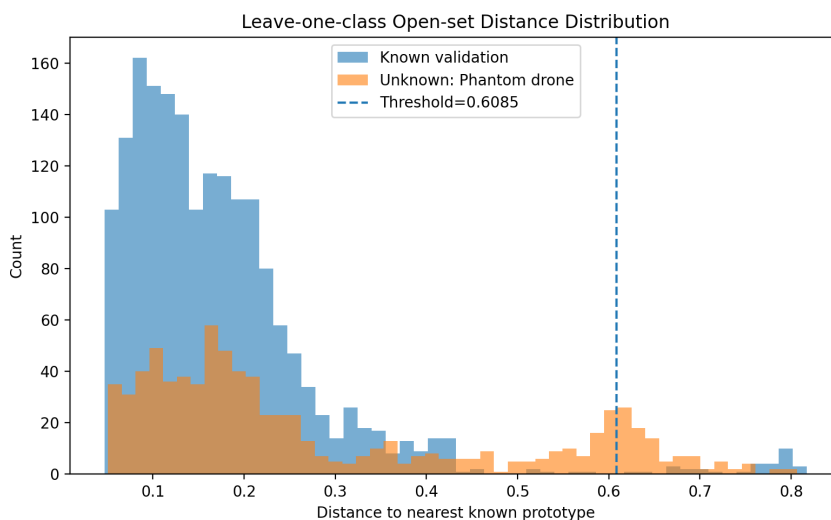


Figure 6. Distance distribution for leave-one-class open-set recognition

图 6. 留一类开集识别距离分布

说明当前特征空间能够较好支持已知类别判别, 但尚未形成稳定的未知目标拒识边界。因此, 开集识别仅作为扩展实验, 后续可结合度量学习和更合理的类别边界建模进一步研究。

5. 结论与展望

5.1. 研究工作总结

本文针对单尺度时频表示难以兼顾不同时间-频率分辨率的问题, 构建 128、256 和 512 三种窗口尺度的 STFT 特征, 并结合 ResNet18 与 SE 通道注意力建立 Multi-STFT ResNet18-SE-Aug 无人机射频分类

模型。实验结果表明, 最终模型在文件级验证集上的 Accuracy 达到 0.7946, Macro F1 达到 0.7833, 整体优于对应的单尺度模型; 在特征域噪声扰动实验中, 多尺度模型在各 SNR 条件下均保持更高识别准确率, 尤其在 10 dB 和 5 dB 条件下优势更明显。结果验证了多尺度时频表示与通道注意力结合用于无人机类型级射频识别的有效性。

5.2. 不足与未来展望

本文实验主要基于 DroneRF 三类无人机数据开展, 数据来源和应用场景仍较有限, 且 Phantom drone 与其他类别之间存在较明显的特征混淆。后续可通过扩展数据集、特征空间分析以及度量学习等方法进一步增强易混淆类别的判别能力。

此外, 本文 SNR 实验采用标准化 STFT 特征域噪声扰动, 尚不能完全等价于真实无线信道环境; 开集实验中的未知目标拒识能力也仍有限。未来可进一步在原始射频采样域研究多径、频偏及无线干扰等复杂条件, 并完善开放集判别方法, 以提高模型在真实低空环境中的泛化与应用能力。

参考文献

- [1] Ezuma, M., Erden, F., Anjinappa, C.K., Ozdemir, O. and Guvenc, I. (2019) Micro-UAV Detection and Classification from RF Fingerprints Using Machine Learning Techniques. 2019 *IEEE Aerospace Conference*, Big Sky, 2-9 March 2019, 1-13. <https://doi.org/10.1109/acro.2019.8741970>
- [2] 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>
- [3] 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 104313. <https://doi.org/10.1016/j.dib.2019.104313>
- [4] Basak, S., Rajendran, S., Pollin, S. and Scheers, B. (2021) Drone Classification from RF Fingerprints Using Deep Residual Nets. 2021 *International Conference on Communication Systems & Networks (COMSNETS)*, Bangalore, 5-9 January 2021, 548-555. <https://doi.org/10.1109/comsnets51098.2021.9352891>
- [5] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [6] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
- [7] Zeng, Y., Gong, Y., Liu, J., Lin, S., Han, Z., Cao, R., et al. (2024) Multi-Channel Attentive Feature Fusion for Radio Frequency Fingerprinting. *IEEE Transactions on Wireless Communications*, **23**, 4243-4254. <https://doi.org/10.1109/twc.2023.3316286>
- [8] Zahid, M.U., Nisar, M.D. and Shah, M.H. (2022) Radio Frequency Fingerprint Extraction Based on Multiscale Approximate Entropy. *Physical Communication*, **55**, Article 101927. <https://doi.org/10.1016/j.phycom.2022.101927>
- [9] Ezuma, M., Erden, F., Kumar Anjinappa, C., Ozdemir, O. and Guvenc, I. (2020) Detection and Classification of UAVs Using RF Fingerprints in the Presence of Wi-Fi and Bluetooth Interference. *IEEE Open Journal of the Communications Society*, **1**, 60-76. <https://doi.org/10.1109/ojcoms.2019.2955889>
- [10] Allen, J.B. and Rabiner, L.R. (1977) A Unified Approach to Short-Time Fourier Analysis and Synthesis. *Proceedings of the IEEE*, **65**, 1558-1564. <https://doi.org/10.1109/proc.1977.10770>
- [11] Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization. arXiv: 1412.6980.