

基于声纹特征的无人机个体识别技术研究

张梦, 乔金兰

引用本文

张梦, 乔金兰. 基于声纹特征的无人机个体识别技术研究[J]. 计算机科学, 2025, 52(11): 444-451.

ZHANG Meng, QIAO Jinlan. [Research on Individual Unmanned Aerial Vehicles Identification Technology Based on Voiceprint Characteristics](#) [J]. Computer Science, 2025, 52(11): 444-451.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[结合动态分析的内存安全漏洞模糊测试方法](#)

Dynamic Analysis Based Fuzz Testing for Memory Safety Vulnerabilities

计算机科学, 2025, 52(11): 382-389. <https://doi.org/10.11896/jsjcx.241000003>

[面向语音助手的窃听攻击与防御研究现状与挑战](#)

Research Status and Challenges of Eavesdropping Attacks and Defenses Targeting Voice Assistants

计算机科学, 2025, 52(11): 364-372. <https://doi.org/10.11896/jsjcx.250300047>

[SCDDA:基于SCA和Dinkelbach的空-天-地一体化网络无人机轨迹与计算卸载优化方法](#)

SCDDA:SCA and Dinkelbach-based Approach for UAV Trajectory and Computation Offloading in Space-Air-Ground Integrated Networks

计算机科学, 2025, 52(11): 270-279. <https://doi.org/10.11896/jsjcx.241100163>

[基于特征增强与上下文融合的无人机小目标检测算法](#)

UAV Small Object Detection Algorithm Based on Feature Enhancement and Context Fusion

计算机科学, 2025, 52(11): 131-140. <https://doi.org/10.11896/jsjcx.241000017>

[格上具有多功能的属性基加密](#)

Multi-functional Attribute Based Encryption from Lattices

计算机科学, 2025, 52(10): 412-422. <https://doi.org/10.11896/jsjcx.240600137>

基于声纹特征的无人机个体识别技术研究

张梦^{1,2} 乔金兰³

1 浙江交通职业技术学院 杭州 311112

2 浙江大学网络空间安全学院 杭州 310058

3 联勤保障部队工程大学勤务指挥系 重庆 401331

(zhangmeng@zjvtit.edu.cn)

摘要 随着人工智能以及通信技术的快速发展,无人机在各行业的应用日益广泛。在低空物流运输领域,无人机凭借高效、便捷和低成本的优势,展现出巨大的应用潜力。然而,合法无人机在执行配送任务时,极易遭受假冒攻击。发货方仅依据外观特征,特别是当恶意第三方采用与合法无人机型号相同的无人机实施假冒行为时,难以准确判断前来取货的无人机是否具备合法性。为了有效解决这一问题,提出一种基于声纹特征的无人机个体身份识别系统 UVBRS。首先,通过移动设备录制无人机悬停时的音频,并利用经验小波变换去除无人机音频信号中的高频噪声,以提高信噪比。然后,基于同一型号不同无人机音频信号的频谱特征设计特制滤波器组,实现对关键特征的精确提取。最后,结合 Open-Max 算法构建长短期记忆网络分类模型,使其能够处理开放集分类问题,进一步提升系统识别能力。实验结果表明,该系统能够以 99.8% 的准确率实现同一型号不同无人机的个体识别,并以 99.5% 的成功率识别出非法无人机,有效抵御假冒攻击。

关键词: 无人机; 声纹; 身份识别; 安全; 物流配送

中图分类号 TP391

Research on Individual Unmanned Aerial Vehicles Identification Technology Based on Voiceprint Characteristics

ZHANG Meng^{1,2} and QIAO Jinlan³

1 Zhejiang Institute of Communications, Hangzhou 311112, China

2 School of Cyber Science and Technology, Zhejiang University, Hangzhou 310058, China

3 Department of Logistics Command, Engineering University of the Joint Logistics Support Force, Chongqing 401331, China

Abstract With the rapid development of artificial intelligence and communication technologies, unmanned aerial vehicles(UAVs) are increasingly being applied across various industries. In the field of low-altitude logistics and transportation, UAVs demonstrate significant potential due to their efficiency, convenience, and low cost. However, legitimate UAVs performing delivery tasks are highly susceptible to spoofing attacks. Shippers relying solely on visual characteristics, especially when malicious third parties use UAVs of the same model as legitimate ones, find it difficult to accurately determine whether an incoming UAV is authorized to pick up a package. To effectively address this issue, an individual UAV identification system based on voiceprint features is proposed. Firstly, audio of a hovering UAV is recorded using a mobile device, and empirical wavelet transform is applied to remove high-frequency noise from the UAV's audio signal, thereby improving the signal-to-noise ratio. Then, a filter bank is designed based on the spectral characteristics of audio signals from different UAVs of the same model, enabling efficient extraction of key audio features. Finally, a long short-term memory network model incorporating the Open-Max algorithm is constructed to handle open-set classification problems, further enhancing the system's recognition capability. Experimental results demonstrate that the proposed system achieves an accuracy of 99.8% in identifying individual UAVs of the same model and a success rate of 99.5% in detecting unauthorized UAVs, effectively mitigating spoofing attacks.

Keywords Unmanned aerial vehicle, Voiceprint, Identity recognition, Security, Logistics and delivery

到稿日期:2025-03-14 返修日期:2025-07-18

基金项目:国家自然科学基金(62372406);浙江交通职业技术学院高层次人才项目(2025rcxm01)

This work was supported by the National Natural Science Foundation of China(62372406) and High-level Talent Project of Zhejiang Institute of Communications(2025rcxm01).

通信作者:乔金兰(qiaojinlan@163.com)

1 引言

近年来,人工智能以及通信技术的快速发展,正以前所未有的速度推动着新一轮的产业革命。无人机以其独特的空中视角、灵活的操作性和广泛的应用场景,在快递配送、环境监测、农业管理等多个领域,都展现出了巨大的潜力和价值。阿里巴巴、京东、美团等电商巨头,纷纷投身其中,全力探索将包裹从仓库直接精准送达客户手中的高效路径。据中国物流与采购联合会航空物流分会统计,2024年我国新开低空物流航线超140条,北京、上海等城市已开通20多条快递、外卖运输航线。有机构测算,到2035年,我国无人机物流产值有望超万亿元,其在物流业的应用市场将以每年20%的速度增长,国内无人机物流配送领域发展前景极为广阔。

然而,无人机在执行配送任务时,极易遭受假冒攻击。当物流公司派出合法无人机到发货方提取包裹时,恶意第三方会使用非法无人机伪装成合法无人机,以此窃取包裹。因此,无人机身份识别技术在判定无人机获取货物合法性方面,发挥着至关重要的作用。目前,有基于雷达、视觉、射频、音频等方法对无人机进行身份识别。

基于雷达的方法利用反射的雷达信号来识别无人机^[1-2]。通过分析不同无人机的微多普勒特征,Yan等^[3]实现了对不同无人机类型(如固定翼、多旋翼和混合垂直起降无人机)的分类。Han等^[4]进一步验证了在无人机体积小或距离较远导致信号衰减的情况下,也可以利用微多普勒特征进行检测。文献^[5]研究了使用毫米波雷达分类3种多旋翼无人机型号的检测方法,准确率为90.10%。虽然基于雷达检测的方法能够提供较大的检测范围以及分类不同类型的无人机,但雷达设备的部署成本高,且无法区分同一型号的不同无人机。

基于射频特征的方法通过拦截无人机与飞行控制模块之间的无线传输信号来实现无人机的识别^[6-8]。文献^[9]提出了一种基于Wi-Fi信号和RF指纹的无人机检测与识别系统。该系统首先进行无人机检测,在检测到无人机后,利用分形维数、轴向积分双谱和平方积分的特性,将其提取为无人机射频指纹。文献^[10]给出了一种基于时频多尺度卷积神经网络的深度学习模型,用于无人机的检测和识别。Swinney等^[11]研究了使用射频指纹来识别无人机型号,在分类10种无人机型号时取得了98%的准确率,并通过实验验证了蓝牙和Wi-Fi信号对分类性能的影响。因此,射频检测识别方法存在的局限是需要额外部署射频通信系统,并且射频信号易受环境干扰的影响。

基于视觉的方法通过摄像头捕捉物体的图像或视频以实现无人机的识别^[12]。Ashraf等^[13]提出了“Dogfight”方法,通过图像处理和目标检测,使用Resnet50模型分类无人机视频中的不同无人机类别。Beher等^[14]利用YOLOv3网络在RGB图像中成功识别并分类无人机,均准确率为74%。Shi等^[15]使用YOLOv4网络识别无人机,其准确率为89.32%。上述文献都只处理了不同型号无人机的分类识别问题,其局限性在于无法准确区分具有相同外观的无人机。

基于声学的方法通过无人机飞行时产生的声音以实现

无人机的识别^[16-18]。文献^[19]和文献^[20]研究了不同型号无人机的分类方法。Anwar等^[19]分别使用梅尔频率倒谱系数(Mel-Frequency Cepstral Coefficients, MFCC)和线性预测倒谱系数(Linear Predictive Cepstral Coefficients, LPCC)提取音频特征,结果表明,使用MFCC方法的分类模型比LPCC方法的准确率更高。Kolamunna等^[20]提出了基于开集的DronePrint系统,通过声纹区分不同型号的无人机,采用长短期记忆网络(Long Short-Term Memory, LSTM)作为分类模型,其分类准确率只有86%。文献^[21]和文献^[22]研究了相同型号无人机的声学特征。Ramesh等^[21]提出了SoundUAV系统,研究了不同型号的无人机电机产生的音频特征,使用MFCC和SVM分类不同型号的无人机电机,分类准确率为99.48%。但在实验过程中,无人机需停放在指定位置并且电机需从无人机上拆卸下来,仅捕捉了电机产生的声音,这限制了其实际运用。Diao等^[22]研究了相同型号不同无人机的分类问题,分别使用MFCC, delta-MFCC和delta-delta MFCC方法提取音频特征,并对比了8种不同的分类模型的性能。但该研究的无人机音频是在室内录音室使用专业麦克风录制的,并且没有对系统抵御安全威胁的能力进行验证分析。

与雷达、视觉以及射频识别等方法相比,基于声学的方法具有诸多优势。它无需额外部署设备,操作过程以及原理较为简单,成本低。在无人机执行配送任务时,由于发货人无法通过外观直接区分无人机的合法性,恶意第三方可能使用与合法无人机型号相同的非法无人机以假冒身份,盗取包裹。由于硬件设备的制造缺陷,其产生的音频信号具有独特性,因而形成了每架无人机独一无二的声纹特征。基于声学的方法能够实现多架相同型号的无人机个体进行区分。

然而,现有基于声学的无人机身份识别的研究存在一定局限。一方面,多数研究只进行了不同型号的无人机身份识别,而没有进行多架同一型号无人机的个体识别。另一方面,多数研究在室内环境进行声音的采集,限制了其实际应用领域的推广。此外,无人机音频易被恶意第三方录制并以重放形式进行假冒攻击,现有研究没有对基于声学的身份识别系统的抗攻击能力进行测试,无法确保系统的安全性。因此,本文提出一种基于声纹特征的无人机个体身份识别系统,以保障无人机物流配送过程的安全运输。主要的创新点如下:

1) 针对无人机音频特性设计精准的特征提取算法。首先利用经验小波变换技术分离并去除信号中的高频噪声,提高无人机音频信号在户外环境下的信噪比。随后通过特制的滤波器组精确提取无人机音频中的基频和谐波分量,这是区分同一型号不同无人机的关键特征。最后从滤波后的信号中利用离散余弦变换技术提取倒频谱系数,并结合信号的能量信息共同构建特征向量。

2) 结合Open-Max算法构建具备开放集识别能力的网络模型。为进一步提高系统识别能力,采用双向长短期记忆(Bidirectional Long Short-Term Memory, Bi-LSTM)网络作为核心分类器并结合Open-Max算法使系统能够识别出未知样本。这一特性对于无人机音频识别系统尤为重要,因为它不仅能够准确认证已注册的无人机,还能有效应对未注册无人机的出现,为系统提供了更高的安全性和灵活性。

3) 身份识别系统的综合性能验证。为了验证本文设计的身份识别系统抵御攻击的能力,在户外环境下进行了多次的实际飞行数据采集工作。通过综合性实验验证了身份识别系统的有效性、鲁棒性以及抵御假冒攻击的安全性能。

2 系统模型

2.1 基于声纹验证的无人机物流配送模型

物流配送公司可以利用无人机提供低空运输服务,从而实现无接触且高效的物流配送。图 1 展示了基于声纹验证的无人机物流配送应用场景。当发货方有物流配送需求时,可以通过指定平台(如物流公司提供的移动应用程序)下订单,填写详细的物流信息,如取件地点、目的地地址及包裹信息等。接到配送订单后,物流公司派遣合法无人机前往指定取件地点。配备导航系统的无人机,能够自动导航到取件点。在取件地点,无人机会在指定位置降落或悬停,发货方可以通过手机录制无人机的声音,并通过身份识别系统验证该无人机身份的合法性。当识别为合法无人机后,发货方便可将包裹装载到无人机上进行配送。装载完毕后,无人机沿预定航线飞往目的地,将货物派送给接收方。

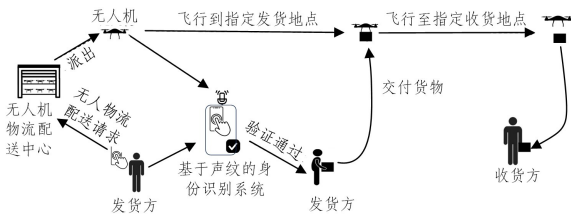


图 1 基于声纹验证的无人机物流配送应用场景

Fig. 1 Application scenarios of UAV logistics and delivery based on voiceprint verification

2.2 威胁模型

在发货阶段,发货方会借助物流配送公司的应用程序发起需要无人机配送业务的请求。随后,物流配送公司便会派遣合法无人机前往发货人的位置取件。在这个过程中,无人机极易遭受假冒攻击。恶意第三方能够通过网络窃取物流信息,进而派遣非法无人机前往取件地点,意图窃取包裹。发货方在看到无人机抵达指定取货地点时,仅从外观上很难判断该无人机是否合法,尤其是当攻击者使用与合法无人机相同型号的无人机进行假冒时。在这种情况下,发送方极有可能将包裹误交给非法无人机,最终致使包裹丢失。冒充攻击模型如图 2 所示。

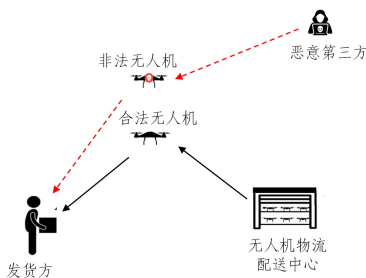


图 2 威胁模型

Fig. 2 Threat model

2.3 无人机音频特征

为了实施基于声纹的无人机身份识别系统以抵御上述攻击,首先需要研究无人机音频的特征,探索其独特的声纹信息。根据空气动力学原理,无人机电机以非常高的速度驱动螺旋桨,通过推动螺旋桨周围的空气向下,从而实现向上动力。无人机声音的产生主要来自其旋转的螺旋桨和运行的电机。由于硬件设备在制造过程中存在缺陷,每架无人机都有其独特的声纹。例如,转子的任何位置出现偏差都会导致电机噪音。此外,不同螺旋桨之间的制造质量差异也会导致不同的风噪音。因此,基于电机和螺旋桨的独特组合,即使是同一型号的无人机,其音频特征也大不相同。

图 3 展示了在相同环境条件下录制的两架大疆 Mini2 无人机悬停时的音频信号频谱信息。

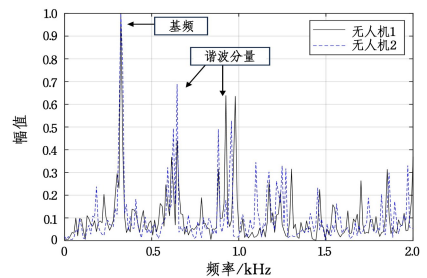


图 3 同一型号不同无人机音频频谱图

Fig. 3 Audio spectrogram of different UAVs of the same model

从图 3 中可以看出,无人机音频主要由基频及其相应的高次谐波组成,而这些谐波在 $0 \sim 8$ kHz 内呈现规则分布,并且这些谐波的频率大致为基频的整数倍。无人机音频信号的基频由其步进电机的 KV 值决定。KV 值表示在无负载时电机的每分钟转速 (Revolutions Per Minute, RPM) 与电池提供的电压的比值。通过将 KV 值与无人机电池电压相乘,可以得到一个 RPM 值,将其除以 60 后即即为以 Hz 为单位的基频。不同型号的无人机由于步进电机的 KV 值不同,会影响其基频及谐波。从图 3 中可以看出,这两架大疆 Mini2 无人机的音频基频都在 320 Hz 左右。由于这两架无人机是同一型号,因此其音频信号的基频信息大致相同,仅在振幅上有轻微差异。另一方面,可以观察到在高次谐波成分,尤其是 $0 \sim 2$ kHz 内,两架大疆 Mini2 无人机的音频谐波分量的振幅和频率存在显著差异,这主要是由硬件制造的细微差异导致的。因此,利用无人机音频的基频以及谐波分量的差异,可以实现对同一型号不同无人机个体的区分。

3 基于声纹的身份识别系统设计

基于不同无人机在音频频谱上的差异,设计基于音频指纹的无人机身份认证系统,具体的技术架构及工作流程如图 4 所示。系统工作流程包括信号预处理、特征提取以及分类识别 3 个阶段。

在信号预处理阶段,为减轻户外环境噪声对无人机音频信号质量的干扰,采用经验小波变换 (Empirical Wavelet Transform, EWT) 方法消除高频噪声,从而增强音频信号的信噪比。在特征提取阶段,首先利用快速傅里叶变换 (Fast Fourier Transform, FFT) 将去噪后的音频信号转换为频域;

然后针对无人机音频的频谱特性,设计专用滤波器组,以实现不同无人机音频特征的精确提取;最后,对滤波后的信号进行离散余弦变换(Discrete Cosine Transform, DCT)获得倒频谱系数并融合信号能量信息构建特征向量。在分类识别阶段,

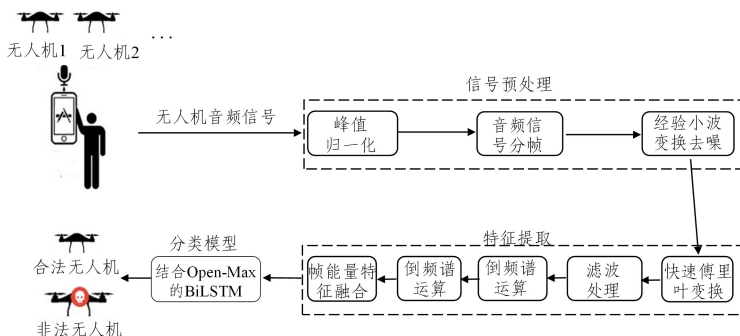


图4 基于声纹的无人机身份识别系统工作流程

Fig. 4 Workflow of a voiceprint-based UAV identification system

3.1 音频信号预处理

在从无人音频信号中提取特征之前,需对原始音频进行预处理,以确保后续信号分析的准确性和稳定性。首先,对音频信号峰值进行归一化,平衡不同录音信号的振幅水平。在录制无人机音频时,即使在噪声最小的环境中,由于操作上的微小差异,移动设备捕捉到的相同无人机音频信号的振幅水平也可能有所不同。通过峰值归一化,可以消除音频信号中的振幅依赖性。归一化后,将每个音频文件分割为较小的帧,以便特征提取。本文选帧的长度为0.1 s,且每个帧与相邻帧有50%的重叠。

录制的音频信号包含各种噪声成分,包括环境噪声和移动设备的仪器噪声,因此,需要对录制的音频信号进行去噪处理。EWT是一种自适应的信号分解方法,其通过分解信号并去除包含噪声的频带成分,相对于普通小波变换方法可以更有效地滤除信号中的噪声,提高信号的质量。其核心思想是从待分析信号的傅里叶频谱中,自适应地构造一组“经验小波”滤波器组,然后利用这些滤波器对信号进行分解,从而获得不同频带上的子信号。EWT的主要步骤如下。

首先对音频信号 $x(t)$ 进行傅里叶变换,得到频域 $X(f)$,计算式为:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \quad (1)$$

然后对 $X(f)$ 进行频谱分析,通过检测信号傅里叶谱中的局部极值来确定频率边界,边界频率点 ω_i 由傅里叶谱的极大值确定,即:

$$\omega_1, \omega_2, \dots, \omega_n = \arg \max |X(f)| \quad (2)$$

这些边界将频域划分为多个带通区域,每个区域对应一个经验小波函数。对于第 n 个频带 $[\omega_n, \omega_{n+1}]$,其对应的小波函数 $\varphi_n(f)$ 可以表示为:

$$\varphi_n(f) = \begin{cases} \sin\left(\frac{\pi(|f| - \omega_n)}{2(\omega_{n+1} - \omega_n)}\right), & \omega_n \leq |f| \leq \omega_{n+1} \\ \cos\left(\frac{\pi(|f| - \omega_{n+1})}{2(\omega_{n+2} - \omega_{n+1})}\right), & \omega_{n+1} \leq |f| \leq \omega_{n+2} \end{cases} \quad (3)$$

通过经验小波函数构建小波滤波器组,使其自适应于特定信号的频谱分布,有效地提取信号特征。然后利用卷积

结合 Open-Max 算法与双向长短期记忆网络(Bi-LSTM),构建了一个适用于开放集分类的模型,从而使系统具备识别未注册无人机的能力。本文在前期工作^[23]的基础上,进一步对去噪方法、滤波器设计进行了优化,提高了无人机的识别性能。

计算不同尺度的小波系数,具体公式为:

$$W_n(t) = \int_{-\infty}^{\infty} x(\tau) \varphi_n^*(\tau - t) d\tau \quad (4)$$

其中, $W_n(t)$ 是对应于第 n 个小波基的系数, $\varphi_n^*(t)$ 是小波函数的共轭,本文采用 db8 小波函数对音频信号进行经验小波变换。进行小波分解后,与噪声相关的小波系数主要存在于高频域,且幅度相对较小。采用软阈值方法消除幅度较小的小波系数,然后利用小波重构获得去噪后的音频信号。

3.2 特征提取

无人机悬停期间其音频信号的频率成分随时间保持相对稳定。因此,利用 FFT 将去噪后的信号转换到频域进行处理。在进行 FFT 变换之后,现有研究一般利用 MFCC 对频谱系数进行转换,以获得音频特征。MFCC 是基于人类听觉系统的工作原理设计的,其利用梅尔滤波器组模拟了人耳对音频信号频率的感知。图5展示了一个包含26个梅尔滤波器的 MFCC 滤波器组。可以看出, MFCC 滤波器组在低频区域包含更多的滤波器,而在高频区域的滤波器较少。

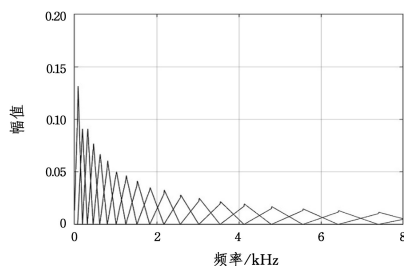


图5 包含26组 Mel 滤波器的 MFCC

Fig. 5 MFCC with 26 Mel filters

然而,同一型号不同无人机的频谱在基频大致相同,它们最大的差异在于高频谐波成分的幅度和频率的不同。因此, MFCC 方法并不适用,必须设计能够符合无人机音频独特特征的特定滤波器组。

因此,本文针对无人机音频特性设计了专用滤波器组(Specialized Audio Filter Bank, SAFB),以精确提取无人机音频的频率特征。根据现有研究,大多数商用无人机的基频通常在150~400 Hz,并且不同无人机的谐波分量在0~2 kHz

的频率范围差异明显,是区分不同无人机的重要特征。因此,在 0~2 kHz 的低频段内,设置第一个滤波器的中心频率 $f_1 = 200$ Hz,并均匀线性分布至 2 kHz,可以表示为:

$$f_k = 200 + (k-1)\Delta \quad (5)$$

$$\Delta = \frac{2000 - 200}{N_{\text{low}} - 1} \quad (6)$$

其中, $k=1, 2, \dots, N_{\text{low}}$, N_{low} 为低频段滤波器的个数。由于在 2 kHz 以上高频范围内的谐波成分的频率和幅度差异相对较小,因此设计了间距随频率增加而非线性变化的三角滤波器。令 $m=1, 2, \dots, N_{\text{high}}$, 则第 $N_{\text{low}} + m$ 个滤波器中心为:

$$f_{N_{\text{low}} + m} = 2000 + \beta m^\rho \quad (7)$$

$$\beta = \frac{F_{\text{max}} - 2000}{(N_{\text{high}})^\rho} \quad (8)$$

对任一滤波器 k , 定义其邻近中心频率为 f_{k-1}, f_k, f_{k+1} , 该滤波器的下边界与上边界为中心的中点, 表示为:

$$l_k = \frac{f_{k-1} + f_k}{2} \quad (9)$$

$$u_k = \frac{f_k + f_{k+1}}{2} \quad (10)$$

其中, $f_0 = 0, f_{k+1} = F_{\text{max}}$, 则三角滤波器组 $H_m(f)$ 可以表示为:

$$H_m(f) = \begin{cases} A_k \cdot \frac{u_k - f}{u_k - f_k}, & f_k < f \leq u_k \\ A_k \cdot \frac{f - l_k}{f_k - l_k}, & l_k < f \leq f_k \\ 0, & f \leq l_k \text{ 或 } f \geq u_k \end{cases} \quad (11)$$

$$A_k = \left(\frac{f_k}{2000} \right)^{-\delta} \quad (12)$$

在设计滤波器组时涉及到的相关参数取值设置及依据如表 1 所列。

表 1 参数总结

Table 1 Parameter summary

描述	参数	值	参考文献
音频采用频率	f_s	44.1 kHz	[21]
采样点数	N	1024	[21]
最大截止频率	F_{max}	8 kHz	[21]
低频滤波器个数	N_{low}	10	[23-25]
高频滤波器个数	N_{high}	16	[23-25]
衰减参数	δ	0.5	[23-25]
中心频率参数	ρ	1.6	[23-25]

根据上述参数值设计的滤波器组如图 6 所示。

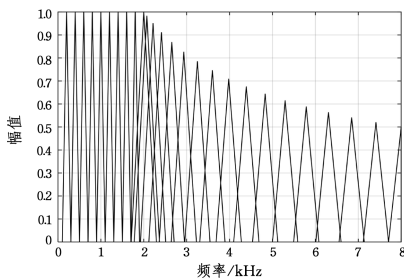


图 6 包含 26 组滤波器的 SAFB

Fig. 6 SAFB with 26 filters

然后,利用 DCT 进行倒频谱运算提取特征系数。DCT 是语音识别和音频分析领域的重要工具,通过 DCT 对滤波后

的信号进行倒频谱变换,能进一步将频谱信息压缩到低维空间中,同时去除频率分量之间的相关性,其计算式为:

$$MFC(\omega, n) = \sqrt{\frac{2}{M} \sum_{m=0}^{M-1} \log[S(i, m)]} \cos\left(\frac{\pi i(2m-1)}{2M}\right) \quad (13)$$

其中:

$$S(i, m) = \sum_{k=0}^{N-1} |X(i, k)|^2 Z_m(k) \quad (14)$$

i 表示帧的数量, $1 \leq \omega \leq W, W=26$ 表示倒谱系数数量。

为了提升后续分类任务的准确性,本文在计算倒谱系数的基础上,进一步计算了每帧的能量总值和均方差值这两个能量要素,并将这些能量要素与倒谱系数合并,共同构成分类模型的 28 维输入特征向量。

3.3 面向开放集的分类模型构建

传统的循环神经网络在处理长序列时,可能会遇到梯度消失或爆炸的问题,导致对长期依赖关系的学习能力受限。为了解决这一问题,长短期记忆网络(LSTM)被引入,它通过设计特殊的门控机制,有效地缓解了梯度问题。然而,标准的 LSTM 只能利用从过去到当前的信息,即进行单向处理。Bi-LSTM 通过在网络中引入两个并行的 LSTM 层,分别是一个正向处理序列和一个反向处理序列,能够同时考虑前后文信息,更加准确地理解音频信息。因此,本文采用 Bi-LSTM 作为分类器。

Open-Max 是一种专为开放集识别任务设计的算法,能有效处理来自未知类别的样本。在传统的识别任务中,模型通常被训练以识别已知类别,并假设所有未知样本都属于这些已知类别之一。然而,在开放集场景中,模型能够在识别已知类别样本的同时,检测和拒绝来自未知类别的样本。由于实际场景中,恶意第三方经常使用未注册的无人机进行攻击,因此将 Bi-LSTM 与 Open-Max 进行结合,以实现未知无人机的识别。首先,使用带标签的数据在 Bi-LSTM 模型上进行闭集训练,以学习特征和模式。在此过程中,从训练后的 Bi-LSTM 模型中收集激活向量(Activation Vector, AV)。然后,这些 AV 被传递到 Open-Max 层进行分析,计算已知类别和未知类别的概率分数。最终,设定一个预定的概率阈值,则以确定样本是属于已知类别还是未知类别。在分类预测阶段,对于每个输入样本,模型检查 Open-Max 输出向量中的概率值。如果任何概率值超过设定的阈值,则模型将样本分类为概率最高的类别。如果所有概率值都低于阈值,模型将样本分类为未知类别。

4 实验结果

4.1 实验设置

为获取清晰的无人机音频数据,选择了相对安静的户外地点,在无人机悬停期间进行音频的录制。无人机在悬停阶段通常处于稳态飞行状态,发动机及旋翼的工况相对稳定,有利于无人机音频信号的采集。具体的录制操作中,实验人员手持移动设备,与无人机保持在长 2 米、高 3 米的位置,以更真实地模拟包裹收发场景。为有效降低天气、温度等环境因素对实验结果的干扰,实验人员每天进行 3 次音频数据的

采集,每次录制每架无人机悬停 3 分钟的音频,并连续进行 7 天数据收集。实验环境如图 7 所示。

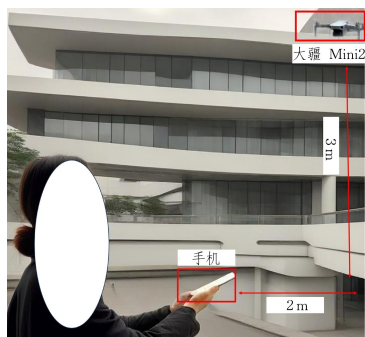


图 7 实验环境

Fig. 7 Experimental environment

音频采集对象包含 6 架大疆 Mini2 无人机、2 架大疆 FPV 无人机以及 2 架 Fama 无人机,共 3 种型号 10 架无人机,如图 8 所示。将采集到的数据按照 15%、15% 和 70% 的比例划分,分别用于预测、验证和训练模型。验证系统从收到无人机音频信息到识别出无人机的身份标识,总共需要 0.1 s。

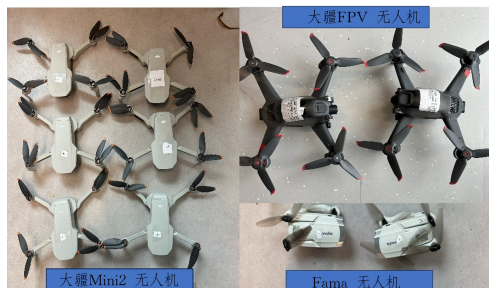


图 8 实验中采用的无人机

Fig. 8 UAVs used in the experiments

4.2 整体性能

为验证本文提出的基于声纹的无人机识别系统(Unmanned aerial vehicles Voiceprint Based Recognition System, UVBRS)的性能,分别对多架不同型号无人机以及多架同一型号无人机进行了身份验证。

首先,利用 UVBRS 对大疆 Mini2、大疆 FPV 以及 Fama,这 3 种不同型号无人机进行分类。实验结果表明,UVBRS 的分类准确率达到 100%。进一步地,利用 UVBRS 对 6 架来自同一型号的大疆 Mini2 无人机进行身份验证。实验结果表明,UVBRS 的认证正确率为 99.8%。

为说明 UVBRS 的优越性,与文献[21]方案进行了对比。文献[21]使用 LSTM 模型对无人机进行分类。为了对比 LSTM 模型与本文给出的 Bi-LSTM 模型的性能平均值是否显著不同,对两组模型分别在准确率、精确率、召回率、F1 值上进行独立样本 t 检验, t 检验结果如表 2 所列。

表 2 两种模型的 t 检验结果

Table 2 Results of t -test of two models

指标	t 值	p 值	是否显著($p < 0.05$)
准确率	34.89	2.80×10^{-7}	显著
精确率	42.17	1.57×10^{-7}	显著
召回率	36.14	2.55×10^{-7}	显著
F1	40.57	1.69×10^{-7}	显著

结果表明,两个模型在准确率($t=34.89, p < 0.001$)上的差异具有统计学显著性。同样,精确率($t=42.17, p < 0.001$)、召回率($t=36.14, p < 0.001$)以及 F1 值($t=40.57, p < 0.001$)也均达到了显著水平($p < 0.05$)。综上所述,Bi-LSTM 模型在各项指标上均显著优于 LSTM 模型,说明其整体性能具有统计学优势。

另外,还对比了采用不同特征提取方法的系统识别性能。文献[22]利用了传统的 MFCC 方法对无人机音频进行提取,将其与本文提出的 SAFB 特征提取方法进行对比,实验结果如图 9 所示。基于 MFCC 方法的系统分类准确率为 93.6%,精确率为 93.7%,召回率为 93.7%,F1 值为 93.7%;基于 SAFB 的特征提取方法的所有评价指标均达到了 99.8%,显著优于文献[22]方案。这是由于 MFCC 适用于人类语音信号的处理,而 SAFB 更能精准针对无人机音频频谱特点来提取特征,从而提高系统识别性能。

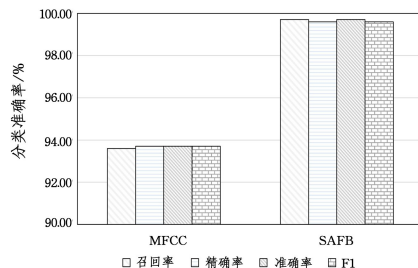


图 9 不同特征提取方法的分类性能对比

Fig. 9 Comparison of classification performance among different feature extraction methods

4.3 安全性能

当发件人通过应用下单后,物流公司会派遣已注册的无人机前来取件。恶意第三方可以通过网络攻击窃取物流信息,并派出一架相同型号的非合法无人机前往取件地址以盗取包裹。在这种情况下,攻击者可以通过合法无人机在执行某一次配送任务的过程中,偷偷录下音频,然后通过用播放器重放录制的无人机音频以实现假冒攻击。

为复现这种攻击场景,首先录制了一架合法的大疆 Mini2 无人机的音频,然后使用 Marshall EmbertonII 蓝牙播放器在另一架无人机悬停时重播录制好的合法无人机音频。实验结果表明,本文设计的身份识别系统以 99.5% 成功率、0.05% 错误拒绝率和 0.05% 的错误接收率识别出无人机的非法身份,有效抵御了假冒攻击,确保货物安全。这是因为,当冒充无人机使用播放器重播合法无人机的音频时,播放器的软硬件设计会自动对音频信号进行一些处理技术,如滤波,从而改变音频的特性;其次,在录音时,移动设备还会捕捉到冒充者自身无人机发出的音频。这些因素使得系统能够区分冒充无人机与合法无人机。

4.4 鲁棒性能分析

4.4.1 不同类型移动设备的影响

由于移动设备的便捷性,用户可能会使用不同类型的手机来录制无人机音频。然而,使用不同类型的手机会导致录音质量的差异。为了测试不同类型移动设备对性能的影响,分别使用华为 Mate20、小米 Note10、OPPO Find X8 和 iPhone

14Pro 这 4 种型号的手机对无人机音频进行了录制,分类性能如图 10 所示。使用 OPPO Find X8 和 iPhone 14Pro 时,系统的分类性能相似,所有评估指标均超过 99.5%;使用 Mate20 和 Note10 时,评估指标分别在 98.5% 和 98.4% 左右。

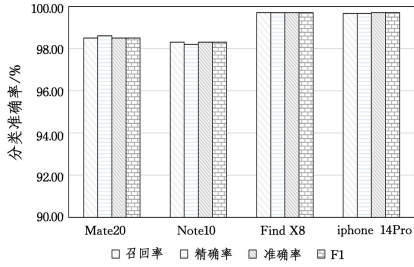


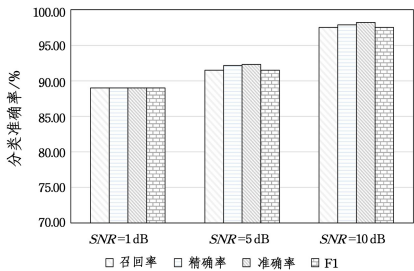
图 10 不同移动设备对验证性能的影响

Fig. 10 Impact of different mobile devices on verification performance

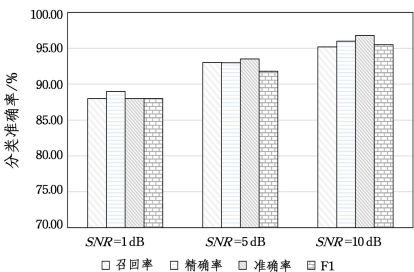
4.4.2 不同录制环境的影响

在室外环境中录制无人机音频时,可能会存在不同程度的噪声。为了评估识别系统在不同环境中的鲁棒性,本文从开放的音频网站下载了 3 种场景的噪声样本,分别为城市室内、城市室外以及郊野室外场景^[26],并将每个噪声样本以 1 dB, 5 dB 和 10 dB 的信噪比 (Signal-to-Noise Ratio, SNR) 分别加入原始无人机音频中。

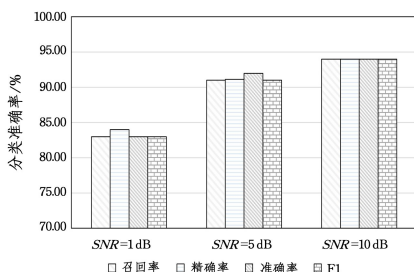
图 11 展示了系统在 3 种场景下不同信噪比下的分类性能表现。



(a) 城市室内



(b) 城市室外



(c) 郊野室外

图 11 不同录制环境对验证性能的影响

Fig. 11 Impact of different scenarios on verification performance

从图中可以看出,3 个不同场景下的分类性能随着信噪比的增加而提高。当 $SNR=1$ dB 时,表明此时噪声和信号的强度非常接近,录制环境极为恶劣。在城市室外和城市室内场景中,所有评估指标均超过了 88%;在郊野室外场景中超过 83%。当 $SNR=5$ dB 时,在 3 个场景中,特别是在郊野室外场景中,识别性能有了显著提升,所有评估指标均超过了 91%。当 $SNR=10$ dB 时,在 3 个场景中的所有评估指标均超过了 94%,在城市室内场景的分类准确率最高达 98.1%。相对于文献[23],由于本文采用了经验小波变换以及调整了滤波器系数,提高了系统去噪能力以及特征提取的准确度,在不同信噪比下,尤其在低信噪比情形下,系统的分类性能显著提升。

4.4.3 不同录制距离的影响

在装载和卸载包裹时,用户与无人机之间的距离变化可能会影响认证性能。图 12 展示了不同距离下的认证性能。在 $s=2$ m, $h=3$ m 以及 $s=3$ m, $h=2$ m 的场景下,所有评估指标均超过了 98.5%。即使在 s 和 h 均为 3 m 的情况下,所有评估指标也均超过了 98%。

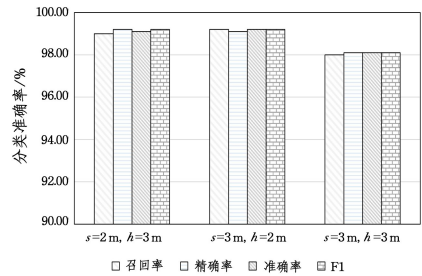


图 12 不同录制距离对验证性能的影响

Fig. 12 Impact of different recording distances on verification performance

结束语 本文提出了一种基于声纹的无人机个体识别系统,通过无人机音频中的基频和谐波成分差异来识别同一型号的不同无人机,以确保无人机配送过程中的货物安全。首先,利用 EWT 去除无人机音频的高频噪声;然后,针对无人机的音频特性设计特定滤波器组用于提取无人机音频频谱特征,并利用 DCT 技术提取倒频谱系数结合能量信息形成特征向量;最后,构建了结合 Open-Max 的 Bi-LSTM 分类模型,使得系统具有开放集识别能力。实验结果表明,提出的基于声纹特征的身份识别系统能够以 99.8% 的准确率区分同一型号的不同无人机。此外,它能够有效抵御基于无人机音频重放的假冒攻击,准确率达到 99.5%,确保货物能交付到合法无人机手中。

然而,目前系统主要面向的是单架无人机进行身份识别的场景,当研究范围扩展至无人机群执行包裹配送任务时,将面临更复杂的问题。在这种情况下,多架无人机产生的重叠音频信号会引入显著干扰,从而对认证过程构成重大挑战。为了解决这一问题,正在探索将独立成分分析与盲源分离技术相结合的方法,将无人机群产生的混合复杂音频进行分离,解开重叠信号,使系统能够在无人机群场景下应用。

参考文献

[1] KUMAWAT H C, CHAKRABORTY M, RAJ A A B, et al. Ar-

- chitecture for Micro-Doppler-Based Small Unmanned Aerial Vehicle(SUAV) Targets' Detection and Classification[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 8504011.
- [2] SUN Y, ABEYWICKRAMA S, JAYASINGHE L, et al. Micro-Doppler signature-based detection, classification, and localization of small UAV with long short-term memory neural network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(8): 6285-6300.
- [3] YAN J, HU H, GONG J, et al. Exploring Radar Micro-Doppler Signatures for Recognition of Drone Types [J]. Drones, 2023, 7(4): 280.
- [4] HAN S K, LEE J H, JUNG Y H. Convolutional Neural Network-Based Drone Detection and Classification Using Overlaid Frequency-Modulated Continuous-Wave Range-Doppler Images [J]. Sensors, 2024, 24(17): 5805.
- [5] ZHANG P, LI G, HUO C, et al. Exploitation of multipath micro-Doppler signatures for drone classification[J]. IET Radar, Sonar & Navigation, 2020, 14(4): 586-592.
- [6] ZHENG Y, ZHANG X, WANG S, et al. Convolutional Neural Network and Ensemble Learning-Based Unmanned Aerial Vehicles Radio Frequency Fingerprinting Identification [J]. Drones, 2024, 8(8): 391.
- [7] HAN J, YU Z, YANG J. Real-world UAV recognition based on radio frequency fingerprinting with Transformer [J/OL]. <https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/emu2.70004>.
- [8] LI Z, CHEN B, CHEN X, et al. Reliable digital forensics in the air: Exploring an RF-based drone identification system [C] // Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2022: 1-25.
- [9] NIE W, HAN Z C, ZHOU M, et al. UAV Detection and Identification Based on WiFi Signal and RF Fingerprint [J]. IEEE Sensors Journal, 2021, 21(12): 13540-13550.
- [10] XIE Y, JIANG P, GU Y, et al. Dual-Source Detection and Identification System Based on UAV Radio Frequency Signal [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-15.
- [11] SWINNEY C J, WOODS J C. RF detection and classification of unmanned aerial vehicles in environments with wireless interference [C] // 2021 International Conference on Unmanned Aircraft Systems(ICUAS). IEEE, 2021: 1494-1498.
- [12] MPOUZIOTAS D, KARVELIS P, STYLIOS C. Advanced Computer Vision Methods for Tracking Wild Birds from Drone Footage [J]. Drones, 2024, 8(6): 259.
- [13] ASHRAF M W, SULTANI W, SHAH M. Dogfight: Detecting drones from drones videos [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 7067-7076.
- [14] BEHERA D K, RAJ A B. Drone detection and classification using deep learning [C] // 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2020: 1012-1016.
- [15] SHI Q, LI J. Objects detection of UAV for anti-UAV based on YOLOv4 [C] // 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT). IEEE, 2020: 1048-1052.
- [16] DONG Q, LIU Y, LIU X. Drone sound detection system based on feature result-level fusion using deep learning [J]. Multimedia Tools and Applications, 2023, 82(1): 149-171.
- [17] AYDIN İ, KIZILAY E. Development of a new light-weight convolutional neural network for acoustic-based amateur drone detection [J]. Applied Acoustics, 2022, 193: 108773.
- [18] BERNARDINI A, MANGIATORDI F, PALLOTTI E, et al. Drone detection by acoustic signature identification [J]. Electronic Imaging, 2017, 29: 60-64.
- [19] ANWAR M Z, KALEEM Z, JAMALIPOUR A. Machine learning inspired sound-based amateur drone detection for public safety applications [J]. IEEE Transactions on Vehicular Technology, 2019, 68(3): 2526-2534.
- [20] KOLAMUNNA H, DAHANAYAKA T, LI J, et al. Droneprint: Acoustic signatures for open-set drone detection and identification with online data [C] // Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2021: 1-31.
- [21] RAMESH S, PATHIER T, HAN J. Sounduav: Towards delivery drone authentication via acoustic noise fingerprinting [C] // Proceedings of the 5th Workshop on Micro Aerial Vehicle Networks, Systems, and Applications. 2019: 27-32.
- [22] DIAO Y, ZHANG Y, ZHAO G, et al. Drone authentication via acoustic fingerprint [C] // Proceedings of the 38th Annual Computer Security Applications Conference. 2022: 658-668.
- [23] ZHANG M, LU L, WU Y, et al. DroneAudioID: A Lightweight Acoustic Fingerprint-Based Drone Authentication System for Secure Drone Delivery [J]. IEEE Transactions on Information Forensics and Security, 2025, 20: 1447-1461.
- [24] MUGISHA S, WASSAJJA M. Review and analysis on digital filter design in digital signal processing [J]. KIU Journal of Science Engineering and Technology, 2024, 3(2): 12-20.
- [25] HAYES B, SHIER J, FAZEKAS G, et al. A review of differentiable digital signal processing for music and speech synthesis [J]. arXiv: 2308. 15422, 2023.
- [26] NOGUEIRA A F R, OLIVEIRA H S, MACHADO J J M, et al. Sound Classification and Processing of Urban Environments: A Systematic Literature Review [J]. Sensors, 2022, 22(22): 8608.



ZHANG Meng, born in 1988, Ph.D, postdoctoral researcher, associate professor. Her main research interests include Internet of Things security and wireless communication security.



QIAO Jinlan, born in 1990, lecturer. Her main research interests include operational research, big data intelligence and AI.