drones

MDPI

Article

Dataset Augmentation and Fractional Frequency Offset Compensation Based Radio Frequency Fingerprint Identification in Drone Communications

Dongming Li [1,*], Zhaorui Wang [1], Yuting Lai [1] and Huafei Shen [2]

[1] School of Cyber Science and Engineering, Southeast University, Nanjing 210000, China; wangzhaorui@seu.edu.cn (Z.W.); ytlai@seu.edu.cn (Y.L.)
[2] Sumec Group Corporation, Nanjing 210031, China; shf@sumec.com.cn
* Correspondence: lidm@seu.edu.cn

Abstract: The open nature of the wireless channel makes the drone communication vulnerable to adverse spoofing attacks, and the radio frequency fingerprint (RFF) identification is promising in effectively safeguarding the access security for drones. Since drones are constantly flying in the three dimensional aerial space, the unique RFF identification problem emerges in drone communication that the effective extraction and identification of RFF suffer from the time-varying channel effects and unavoidable jitterings due to the constant flight. To tackle this issue, we propose augmenting the training RFF dataset by regenerating the drone channel characteristics and compensate the fractional frequency offset. The proposed method estimates the Rician K value of the channel and curve-fits the statistical distribution, the Rician channels are regenerated using the sinusoidal superposition method. Then, a probabilistic switching channel is also set up to introduce the Rayleigh channel effects into the training dataset. The proposed method effectively addresses the unilateral channel effects in the training dataset and achieves the balanced channel effect distribution. Consequently, the pre-trained model can extract channel-robust RFF features in drone air-ground channels. In addition, by compensating the fractional frequency offset, the proposed method removes the unstable frequency components and retains the stable integer frequency offset. Then, the stable frequency offset features that are robust to environmental changes can be extracted. The proposed method achieves an average classification accuracy of 97% under spatial and temporal varying conditions.

Keywords: drone communications; physical-layer security; radio frequency fingerprint

Citation: Li, D.; Wang, Z.; Lai, Y.; Shen, H. Dataset Augmentation and Fractional Frequency Offset Compensation Based Radio Frequency Fingerprint Identification in Drone Communications. *Drones* **2024**, *8*, 569. https://doi.org/10.3390/drones8100569

Academic Editor: Shiva Raj Pokhrel

Received: 6 August 2024
Revised: 1 October 2024
Accepted: 7 October 2024
Published: 10 October 2024

1. Introduction

In recent years, the civilian drone market has experienced rapid growth, making effective drone identity authentication processes crucial for drone regulation. Compared to more mature wireless solutions for other traditional internet of things (IoT) devices, such as Wi-Fi, Bluetooth, and ZigBee, drone wireless network communication typically provides fewer built-in security modules due to platform energy constraints. This makes their wireless air interfaces highly susceptible to malicious attacks from third parties. As a result, the effective authentication of drones from the ground stations plays key roles in drone communications.

1.1. Physical-Layer Radio Frequency Fingerprinting

The traditional high-layer identity authentication is based on the ciphers for identifying the legacy of unknown nodes. This process often assumes a trusted initial key distribution phase and typically only guarantees computational security, which requires a certain amount of computing power and time to crack the ciphers, and both parties' keys need to be updated regularly. However, drone platforms have the common characteristic of extremely scarce energy and computing resources in mobile embedded systems, and their

limited energy is mainly allocated to motors, Micro Controller Units (MCU), and radio frequency units. MCU is generally responsible for real-time processing tasks, such as image recognition and route planning, and it is difficult to bear the additional computational overhead brought by high-layer security protocols. Limited by the openness and flexibility of wireless networks, traditional high-layer security protocols are relatively vulnerable [1].

For the vulnerability of identity authentication in high-layer network protocols, physical-layer security has received considerable attentions in recent years. As an alternative to the vulnerability of identity authentication processes in higher-layer network protocols to man-in-the-middle attacks, Radio Frequency fingerprinting (RFF) technology has recently gained popularity as a physical layer identity authentication technique. RFF, derived from the inherent tolerances of internal components within the transmitter, is difficult to clone. Identity authentication technology based on RFF can provide enhanced security during the access phase of drone communications. RFF is advantageous in resource constrained platforms such as drones, reducing reliance on computing resources and providing stronger wireless security. Considering the unique operational environments and market demands of drones, promoting the application of RFF technology on drone platforms is of significant practical importance.

As a novel physical layer identity authentication method, RFF was first proposed by Canadian scholars Hall et al. [2]. The core of this technology is to extract unique hardware features generated by manufacturing errors or aging factors within the transmitter, including frequency offset and phase noise of crystal oscillators, DC offset of digital to analog converters, IQ offset of mixers, the nonlinear characteristics of power amplifiers and the frequency response errors of various filters. RFF technology aims to extract these stable features from the transmitter hardware circuit at the receiving end. The hardware features of the receiver will also affect the wireless signal, and the extracted RFF features are the combination of hardware features of the transmitter and receiver. Generally, candidate features that can be used as RFF should have the following properties:

(1) Uniqueness. Uniqueness is the most fundamental property that ensures the use of RFF for identity authentication. Different manufacturers usually extract RFF with high discriminability due to differences in manufacturing processes or procedures. Equipments from the same manufacturer are of small hardware differences, but generally have distinguishability [3].
(2) Universality. Although there are various hardware and signal modulation methods available in the current market, signals emitted from all devices should have extractable RFF features.
(3) Robustness. Factors including temperature, humidity, voltage, channel environment, distance between transmitting and receiving devices, and signal polarization can affect the effectiveness of RFF. The RFF system must remain stable under various factors and have robustness.
(4) Short term invariance. Hardware aging can cause changes in hardware tolerances, resulting in RFF extracted each time in the database have certain differences, and it is necessary to ensure that the extracted RFF remain unchanged for a period of time.

1.2. Literature Review

The unpermitted flying of drones may pose a significant threat to military battlefields or social infrastructure. Therefore, ground stations pose great importance to the identification and early warning of illegal drones in sensitive airspace. For the above application scenarios, some researchers have attempted to use image or audio data to achieve the goal of detecting and recognizing drones. Chalmers et al. [4] proposed that within the restricted airspace of civil airports, high-resolution panoramic cameras can be installed to achieve global perception of the area and detect and identify micro air vehicles, including drones. This method mainly targets at the unpermitted flying events of small or micro aircraft. Researchers have constructed a drone video dataset captured from a special perspective and

combined it with convolutional neural networks and optical flow localization technology to analyze object motion patterns.

Yoo et al. [5] attempted to extract features from the sound waves generated by the high-frequency rotation of motors on quadcopters. Researchers constructed audio sample datasets for over ten unmanned institutions and used LeNet neural network models for classification. The classification accuracy of over ten types of drones reached 95% in the experimental environment. Similar to image classification, achieving drone classification by analyzing the audio information generated by the high-frequency rotation of motors also faces many difficulties in real-world applications. First, there is a strong attenuation in the propagation of sound waves, and some environmental background sounds can mask the sound emitted by the drone itself. Second, such methods will also face the same challenges as drone image classification, namely poor classification performance for the same model of drone.

The above two methods of using drone image and sound features to classify drones will face great instability in practical applications and cannot meet the requirements of high accuracy and reliability for classifying drones of the same model. The uniqueness of RFF makes it possible to distinguish devices of the same model, and classifying drones through RFF has broad prospects. Ground station can extract device RFF and compare them with the stored RFF in the backend database, achieving identification and early warning of illegal intrusion of drones.

In addition, the Micro Air Vehicle Link (MAVLink) is currently the most commonly used protocol for communication between drones and ground stations, as well as between drones. Kunze et al. [6] have demonstrated that the security of the MAVLink protocol is currently at a low level, indicating that the high real-time and lightweight characteristics of the MAVLink protocol are achieved by sacrificing communication security. Researchers used universal software radio peripheral as a third-party ground station, and through reverse analysis of the application layer MAVLink protocol and physical layer frame format, ultimately achieved eavesdropping and drone hijacking of the MAVLink communication link between drones and ground station at a lower cost.

1.3. Contributions and Novelty

Current research on RFF primarily focuses on low-mobility IoT devices, with experiments typically conducted in indoor stationary or low-speed scenarios. Few studies have addressed RFF extraction on high-speed platforms such as drones. In RFF identification, channel fading and device frequency offset are critical factors affecting identification and authentication accuracy. The RFF identification for drone platforms faces more severe challenges due to the time-varying channel fadings. The overall frequency offset of drones is also more complex, typically involving additional Doppler shifts and frequency fluctuations caused by factors such as body vibrations, as well as oscillator frequency offsets. Therefore, designing robust RFF extraction methods that can avoid wireless channel fadings and accurately compensate drone frequency offsets is crucial for RFF in drone communications. Different from the existing studies that mainly focus on RFF on indoor stationary or low-speed nodes, this work studies the drone RFF technology taking high mobility into account. To avoid the adverse strong time-varying channel effects, a data augmentation method based on air-ground channel reconstruction is proposed, enhancing the robustness on drone flight state datasets by improving the distribution of channel effects in the training dataset. To mitigate the adverse impacts of drone frequency offset on RFF extraction, a fractional frequency offset compensation method is proposed, effectively suppressing the fractional frequency offset while reasonably preserving the integer frequency offset. To the best of the authors' knowledge, this is the first work that studies the drone RFF by proposing channel regeneration data augmentation and fractional frequency offset compensation. The main contributions of this work are as follows:

(1) Proposed a Data Augmentation Method Based on Drone Air-Ground Channel Generation: This method first measures the Rician fading factor in the target deployment

environment using the method of moments and fits a probability density curve to the distribution obtained. This curve is then used to randomly generate Rician factors in the data augmentation process, simulating Rician channels based on the sinusoidal superposition method. A probabilistic switching channel is also set up to introduce some Rayleigh channel effects into the training set. This method effectively addresses the overly uniform channel effects in the training set, achieving a balanced channel effect distribution. Consequently, the pre-trained model can extract channel-robust RFF features in drone air-ground channels. This data augmentation method simplifies the registration process for drones in the RFF system, completing registration in about two minutes in an indoor experimental environment.

(2) Proposed a Fractional Frequency Offset Compensation Method for Drones: This paper analyzes the composition of the overall frequency offset in drones, which includes oscillator frequency offset, flight Doppler shift, and frequency fluctuations caused by airframe vibrations. By compensating for the fractional frequency offset, the method removes the easily affected part of the frequency offset while retaining the stable integer frequency offset over long periods. This allows the neural network's feature extraction layer to learn stable frequency offset features that are robust to environmental changes. Compared to traditional methods that either retain or completely remove frequency offset features, this method achieved an average classification accuracy of 97% in stationary state tests under varying spatiotemporal conditions.

(3) Experimental Design Incorporating Drone Mobility: The experiments in this thesis consider not only the stationary state, as in traditional IoT RFF extraction research, but also the hovering and moving states of drones. The data collection for these test environments spans significant temporal and spatial dimensions, effectively assessing the stability and reliability of the proposed methods. A deep neural network model incorporating residual blocks and channel attention mechanisms is designed for feature extraction and classification of drone signals. The use of residual blocks alleviates the problems of gradient vanishing and explosion during training, while the channel attention mechanism enhances the model's focus on transmitter hardware characteristics. Additionally, to verify the importance of the K-factor generation strategy in the proposed data augmentation method, comparisons with the typical uniform distribution K-factor generation strategy were conducted. For the proposed frequency offset handling method, comparisons with two common traditional methods—complete retention and complete removal of frequency offset—were also made. Experimental results show that the proposed RFF extraction method for drone platforms improves classification accuracy by up to 33% in drone flight states (speed 8–12 m/s) compared to noise-only augmentation methods.

2. System Model

The Sikradio radio frequency (RF) module, as a low-power communication module, uses a communication protocol belonging to IEEE 802.15.4 g. The detailed physical layer signal format is shown in Figure 1. The preamble field used for frame synchronization is the part used for RFF extraction. The preamble field provides a good prerequisite for subsequent RFF extraction due to the transmission of fixed and regular data bits.

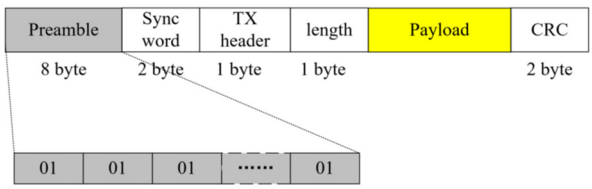

Figure 1. The physical-layer signal format of drones.

The extraction of the preamble sequence serves as the subsequent RFF extraction. The goal of the preprocessing stage is to extract the preamble, including the steps of signal

frame detection, synchronization, and amplitude normalization. For the detection of the starting point of a single frame signal from a drone, a sliding window variance calculation method is adopted while considering time complexity. This method generates two variance peaks at the beginning and end of a single frame signal, and the tail variance peak needs to be discarded. Completing this step is equivalent to distinguishing the effective signal segment from the noise segment and marking the starting position of each frame's effective signal. Then, it is necessary to use the repeated structure of the preamble sequence from the starting position to search for the precise offset of the preamble sequence starting point compared to the signal frame starting point.

Due to the cross-correlation synchronization operation between the received signal and the local sequence, it is usually subject to interference from frequency offset and noise. The preamble sequence specified for PX4 drone 802.15.4 g has a good repetitive structure, thus the timing synchronization of data frames can adopt an autocorrelation synchronization strategy that is insensitive to noise and frequency offset [7]. The following introduces the principle of autocorrelation synchronization. Autocorrelation operation can be considered as the cross-correlation operation of a signal sequence at two different time observation points. The more similar the signal sequences at these two time points, the greater the correlation value. Denote the correlation between two adjacent symbols of length L, i.e., L samples, in the n-th block as $R(n, L, \delta)$, and δ is the offset used to search for the peak, we have

$$R(n, L, \delta) = \sum_{i=0}^{L-1} x(\delta + nL + i)x^*(\delta + L + nL + i) \tag{1}$$

where x denotes the samples.

Denoting the number of symbol repetitions as m in the leading sequence, calculate the autocorrelation sum of adjacent sampling blocks and search for the peak offset δ. Then, the autocorrelation peak is calculated as

$$R_{peak}(\delta) = \sum_{n=0}^{m-1} R(n, L, \delta) \tag{2}$$

If the peak is successfully found, it will be denoted as 1, representing the starting point of the leading sequence. After successfully capturing the starting point of the preamble sequence, it is necessary to normalize the amplitude of the preamble sequence to facilitate subsequent processing. For each complex sampling point $p_{\mathrm{i}} = a + bj$, $I = 1, 2, ..., n$ in the leading sequence, using the maximum amplitude normalization method, the normalized leading sequence p_i' is

$$p_i' = p_i / \max(|p_i|), i = 1, 2, \cdots, n \tag{3}$$

3. The Proposed Scheme

3.1. Estimating Rician K Factor for Drone Channel

The estimation of the Rician K factor of the received signal is performed using the second-order moment estimation method [8]. This method is mainly based on the mathematical relationship between various order distances. First, the parameter expressions of second-order and fourth-order distances are derived using probability distribution functions. Then, the expression of K factor with respect to the order distance is obtained by solving the two equations simultaneously. Then, the estimated value of K factor can be obtained by calculating the second-order and fourth-order distances of the received signal envelope.

Based on the probability density function (PDF) $p_{\xi}(x)$ of the received signal envelope during the Rayleigh fading process, the expressions for each order distance can be obtained as:

$$\begin{aligned} \mu_n &= E[x^n] = \int_0^{+\infty} x^n p_{\xi}(x) dx \\ &= (\sigma^2)^{\frac{n}{2}} \mathrm{e}^{-K} \Gamma\left(\frac{n}{2}+1\right) {}_1F_1\left(\frac{n}{2}+1;1;K\right) \end{aligned} \tag{4}$$

where ${}_1F_1(\cdot;\cdot;\cdot)$ is a convergent hypergeometric function, and $\Gamma(\cdot)$ is a gamma function. According to (4), it is evident that each step distance is related to the unknown parameters K and σ. Thus, to solve the parameter K, at least two different order values need to be calculated. Considering computational complexity, we will use second-order and fourth-order moments to estimate the value of K. The specific second and fourth order distances are

$$\begin{aligned} \mu_2 = E[x^2] = 2\sigma^2 + \alpha^2 = 2\alpha^2(1+K) \\ \mu_4 = E[x^4] = 8\sigma^4 + 8\sigma^2\alpha^2 + \alpha^4 = 4\alpha^4(2+4K+K^2) \end{aligned} \tag{5}$$

Define $f_{n,m}(K) = \frac{\mu_n^m}{\mu_m^n}$, and based on (5), we have

$$f_{2,4}(K) = \left(\frac{\mu_2^2}{\mu_4}\right)^2 = \left(\frac{(1+K)^2}{(2+4K+K^2)}\right)^2 \tag{6}$$

based on which we have

$$\widehat{K} = \frac{-2\widehat{\mu}_2^2 + \widehat{\mu}_4 - \widehat{\mu}_2\sqrt{2\widehat{\mu}_2^2 - \widehat{\mu}_4}}{\widehat{\mu}_2^2 - \widehat{\mu}_4} \tag{7}$$

The estimated values of the second and fourth order distances obtained from (7) can be directly estimated from the received signal envelope, that is

$$\begin{aligned} \widehat{\mu}_2 = E\left[x^2\right] = \frac{1}{N}\sum_{i=0}^{N-1} x_i^2 \\ \widehat{\mu}_4 = E\left[x^4\right] = \frac{1}{N}\sum_{i=0}^{N-1} x_i^4 \end{aligned} \tag{8}$$

The specific K value estimation process is formulated as shown in Algorithm 1. Finally, it is necessary to fit the distribution of K obtained during the aforementioned drone uniform flight to obtain the final probability density curve, and use this curve to drive the random generation of the Rayleigh channel in the data augmentation module.

Algorithm 1. Estimating Rician factor based on the second and fourth order moments.

Input: The preamble matrix $\mathrm{P_b}[N]$, which is the preamble sequence of length N.
Output:K
while$i$=1:N do
$sum_1 = \mathrm{P_b}[i]^2 + sum_1$
$sum_2 = \mathrm{P_b}[i]^4 + sum_2$
end while
$\widehat{\mu}_2 = \frac{sum_1}{N}$
$\widehat{\mu}_4 = \frac{sum_2}{N}$
Calculate $\widehat{K}$ according to (7).
Calculate K=10log$_{10}\widehat{K}$.

3.2. Training Dataset Augmentation

The drone training set enhancement module is shown in Figure 2. This module "Switch 1" is used to select whether to perform this enhancement and retain a certain

proportion of static state samples in the final enhanced training set. "Switch 2" is used to select different fading effect models. The different switching branches are mainly used to ensure that the signal samples of each category in the final enhanced training dataset are relatively uniform.

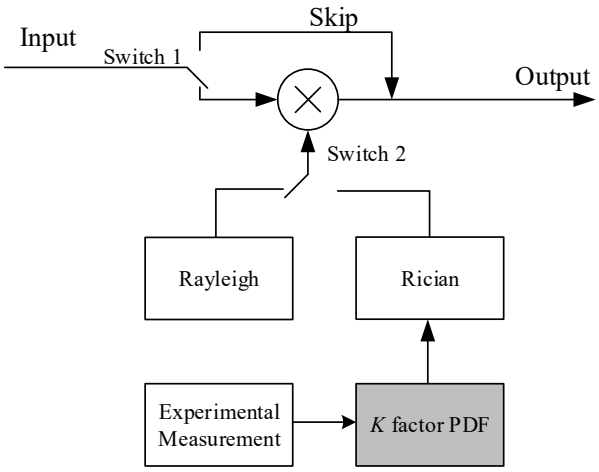

Figure 2. Drone training dataset augmentation module.

Both Rayleigh and Rician models require at least two colored Gaussian random processes $I(t)$ and $Q(t)$, which can be generated using filtering or sine wave superposition methods. The filtering method requires Doppler filtering in the frequency domain and then converts back to the time domain using the inverse fast Fourier transform. The sine wave superposition method is used to obtain a finite number of deterministic sine waves, and the computational complexity of this simulation method is relatively small. Therefore, the data augmentation module uses the sine wave superposition method to generate colored Gaussian random processes. Then, we have

$$\begin{aligned} h_{NLOS}(t) &= I(\mathrm{t}) + jQ(t) \\ I(\mathrm{t}) &= \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \cos(w_d t \cos \alpha_n + \phi_n) \\ Q(\mathrm{t}) &= \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \sin(w_d t \cos \alpha_n + \phi_n) \\ \alpha_n &= \frac{2\pi n + \phi_n}{N} \end{aligned} \tag{9}$$

where n = 1, 2, ..., N, and N is the number of sine waves, w_d is the arc extension caused by the maximum Doppler frequency spread, and α_n and ϕ_n are statistically independent and satisfy a uniform distribution on $[-\pi,\pi)$. Adding LOS path component, we have

$$\begin{aligned} h_{Rician}(t) &= I'(t) + jQ'(t) \\ I'(t) &= \sqrt{\tfrac{K}{K+1}} \cos(w_d t \cos \alpha_0 + \phi_0) + \sqrt{\tfrac{1}{K+1}} I(\mathrm{t}) \\ Q'(t) &= \sqrt{\tfrac{K}{K+1}} \sin(w_d t \cos \alpha_0 + \phi_0) + \sqrt{\tfrac{1}{K+1}} Q(\mathrm{t}) \end{aligned} \tag{10}$$

The final $h_{\mathrm{Rician}}(\mathrm{t})$ and $h_{\mathrm{NLOS}}(\mathrm{t})$ respectively form the Rician and Rayleigh model in the enhancement module shown in Figure 2, and the stationary drone signal samples are enhanced.

3.3. Drone Fractional Frequency Offset Compensation

In practical RF transceiver systems, the carrier frequency offset estimation is divided into two parts: fractional and integer components. This part estimates and compensates the fractional component. The performance of frequency offset estimation is generally

represented by mean square error (MSE). If the actual frequency offset is Δf and the estimated frequency offset is $\Delta \hat{f}$, then MSE is expressed as

$$MSE = \frac{1}{N}\sum_{i=0}^{N-1}\left(\Delta f - \Delta \hat{f}_i\right)^2 \tag{11}$$

For ease of use, the fractional multiple of the oscillator frequency offset is denoted as Δf_{c_f} and the integer multiple is denoted as Δf_{c_i}. Then, the overall frequency offset can be represented as shown in Figure 3, where $f_d(t)$ and $f_v(t)$ are the frequency shifts induced by the Doppler effects and the drone body vibration. If the estimated range of fractional frequency offset is greater than the maximum variation range of device frequency offset, compensation for the overall frequency offset variation can be achieved, while leaving a stable integer frequency offset.

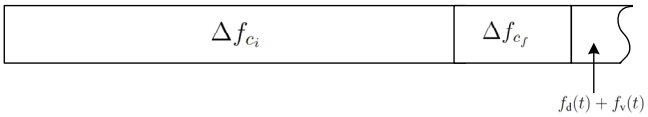

Figure 3. The frequency offset components for drones.

We first use two consecutive repeated symbols in the time domain for correlation operation to achieve rough estimation of fractional frequency deviation. After frequency compensation, we propose to use two repeated blocks with longer time intervals to perform frequency fine synchronization. When there is a frequency offset Δf and a phase offset θ, the two symbol intervals used for frequency offset estimation in the time domain are N sampling points, which can be expressed as

$$\begin{aligned} r_1(n) &= r(n) = x_1(n)e^{j(2\pi\Delta f n T_s+\theta)} \; n = 1, 2, \ldots, L \\ r_2(n) &= r(n+N) = x_2(n)e^{j(2\pi\Delta f (n+N) T_s+\theta)} \; n = 1, 2, \ldots, L \end{aligned} \tag{12}$$

The correlation between two symbol sampling blocks in the time domain is calculated as

$$\begin{aligned} R(N,L) &= \sum_{n=0}^{L-1} r_2(n) r_1^*(n) \\ &= \sum_{n=0}^{L-1} r(n+N) r^*(n) \\ &= \sum_{n=0}^{L-1} x_2(n) e^{j(2\pi\Delta f (n+N) T_s+\theta)} x_1^*(n) e^{-j(2\pi\Delta f n T_s+\theta)} \\ &= e^{j2\pi\Delta f K T_s} \sum_{n=0}^{L-1} x_2(n) x_1^*(n) \\ &= e^{j2\pi\Delta f N T_s} \sum_{n=0}^{L-1} \left| x_1(n) \right|^2 \end{aligned} \tag{13}$$

where * represents conjugation operation, and the phase change caused by frequency offset is angle($R(N,L)$) = $2\pi\Delta f N T_s$. Therefore, a rough estimate of the fractional harmonic offset Δf_1 can be obtained as.

$$\Delta f_1 = \frac{angle(R(N,L))}{2\pi N T_s} \tag{14}$$

and where angle($\cdot$) is used for angle calculation, which can maximize the range of fractional frequency offset estimation when the time interval distance between repeated sampling blocks is minimized. Figure 4 shows the range of fractional frequency offset estimation at different time intervals, where S1, S2, ..., SN are the leading repeated symbols. When using S1 and S2 for fractional multiplication bias estimation, due to the shortest interval between them, it is obvious that they have the largest estimation range. When estimating the sampling blocks composed of S3 and S4, as well as the sampling blocks composed of S5

and S6, the distance between them becomes twice that of the adjacent symbol sampling blocks, resulting in a reduction of the final estimation range to half of the original range. Due to the range of the phase symbol angle(·) of $[-\pi,\pi)$, if adjacent symbol sampling blocks are used for estimation, the maximum range of fractional frequency offset estimation is $|\Delta f_1| \leq 1/2NT_s$.

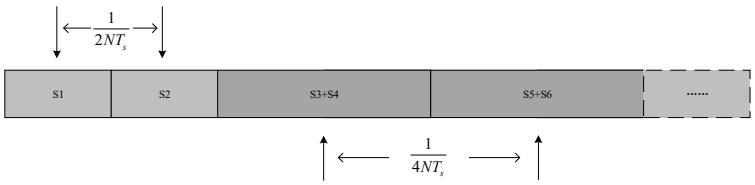

Figure 4. The fractional frequency offset estimation range under different estimation intervals.

For the compensation of Δf_1 obtained from the coarse frequency offset estimation, we use an upward compensation strategy. The compensation process is shown in Figure 5, which compensates for the fractional multiple bias upwards to align the overall frequency bias of drones to the nearest integer multiple. Denoting the sampling rate as f_s, satisfying $f_s = 1/T_s$, N is the estimation interval, which is the number of sampling points between the two symbols in the time domain. Compensating for fractional frequency deviation upwards means rounding up the integer frequency deviation, the final $\Delta f'_1$ to be compensated is

$$\begin{aligned} \Delta f_1' &= \begin{cases} \Delta f_1 - \frac{1}{Nf_s}, & \Delta f_1 > 0 \\ \Delta f_1, & \Delta f_1 < 0 \end{cases} \\ \Delta f_1' &= \begin{cases} \frac{angle(R(N,L))-2\pi}{2\pi NT_s}, & R(N,L) > 0 \\ \frac{angle(R(N,L))}{2\pi NT_s}{}_1', & R(N,L) < 0 \end{cases} \end{aligned} \tag{15}$$

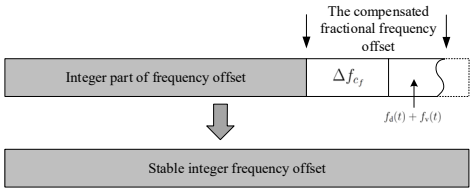

Figure 5. Illustration of upwards compensation of fractional frequency offset.

Therefore, the compensated preamble sequence is

$$\begin{aligned} r'(n) &= r(n)e^{-j2\pi\overset{1}{\Delta} f_1' \frac{n}{f_s}} \\ &= r(n)e^{-j2\pi\Delta f_1' nT_s} \\ &= x(n)e^{j2\pi(\Delta f - \Delta f_{1'})nT_s} \end{aligned} \tag{16}$$

Based on the coarse frequency offset compensation, the residual fractional frequency offset is already very small. Then, it is necessary to select a larger estimation interval N' and a larger repeated sampling block length L' for more accurate fine frequency offset estimation and compensation, which can be expressed as

$$f_2 = \frac{angle(R(N', L'))}{2\pi N' T_s} \tag{17}$$

Similarly, compensating $r'(n)$ yields

$$\begin{aligned} r''(n) &= r'(n)e^{-j2\pi\Delta f_2 nT_s} \\ &= x(n)e^{j2\pi(\Delta f - \Delta f_1' - \Delta f_2)nT_s} \end{aligned} \tag{18}$$

Due to the large time interval selected for the second fine multiple estimation and the fact that the leading sequence with coarse frequency offset compensation already has a very small fractional multiple offset, the accuracy of this estimation is relatively high. Using the method in [9] to analyze the variance of frequency offset estimation, we have

$$\begin{aligned} var(\Delta f_2) &= \left(\frac{L}{2\pi NB}\right)^2 \frac{\left(\sum_{k=0}^{L-1} n'(k)(r'(k+L))^*\right)^2}{\left(\sum_{k=0}^{L-1} r'(k)r'(k+L)^*\right)^2} \\ &+ \frac{\left(\sum_{k=0}^{L-1} r'(k)(n'(k+L))^*\right)^2}{\left(\sum_{k=0}^{L-1} r'(k)r'(k+L)\right)^2} \\ &+ \frac{\left(\sum_{k=0}^{L-1} n'(k)(n'(k+L))^*\right)^2}{\left(\sum_{k=0}^{L-1} r'(k)r'(k+L)^*\right)^2} \\ &= \left[\frac{N}{2\pi L}\right]^2 \frac{(1/2\sigma_s^2)\sigma_n^2 + BK\sigma_s^2(1/2\sigma_n^2) + BK\sigma_n^2(1/2\sigma_n^2)}{(L\sigma_s^2)^2} \\ &\approx \frac{2}{\pi^2 \cdot L \cdot SNR} + \frac{1}{\pi^2 \cdot L \cdot (SNR)^2} \end{aligned} \tag{19}$$

where $n'(k)$ is the additive noise, and where $E\left[|x'(k)|^2\right] = \sigma_s^2$, $E\left[|n'(k)|^2\right] = \sigma_n^2$, $SNR = \sigma_s^2/(2\sigma_n^2)$.

4. Experimental Results and Discussions

To verify the effectiveness of the proposed method, indoor drone stationary scenarios, outdoor drone hovering scenarios, and outdoor drone flight scenarios were evaluated. The indoor drone stationary scene includes five different spatiotemporal environmental datasets, mainly examining the effectiveness and reliability of the proposed method for processing fractional multiples of drones. The outdoor scene includes four different spatiotemporal environmental datasets, including drone hovering, takeoff and landing, low-speed flight (2~4 m/s), and high-speed flight (8~12 m/s). The main focus is to examine the effectiveness of the proposed data augmentation method. In addition, the sampling rate used for drone signal reception is 1.6 MHz. The total length of the preamble sequence is 1600 sampling points.

4.1. Experimental Settings

The specific collection locations of each dataset are shown in Figure 6. There are five sets of signal frames in a stationary state. Among them, the indoor stationary datasets S1, S2, and S3 were collected in Office 1, and S4 is located in the hallway at the entrance of Office 1, S5 is collected in Office 2. The number of data frames for each device in S1, S2, S3, and S4 datasets is equal, while the number of data frames for each device in S5 is approximately equal. Generally, the sending rate of PX4 drones to ground stations is about 700 per minute, the longest collection process for these stationary datasets is about 7–8 min. Then, the collection dates of these five datasets span over several months, which can effectively examine the stability and reliability of the extracted RFF. The specific information regarding the number of data frames, collection date, and transceiver distance is shown in Table 1.

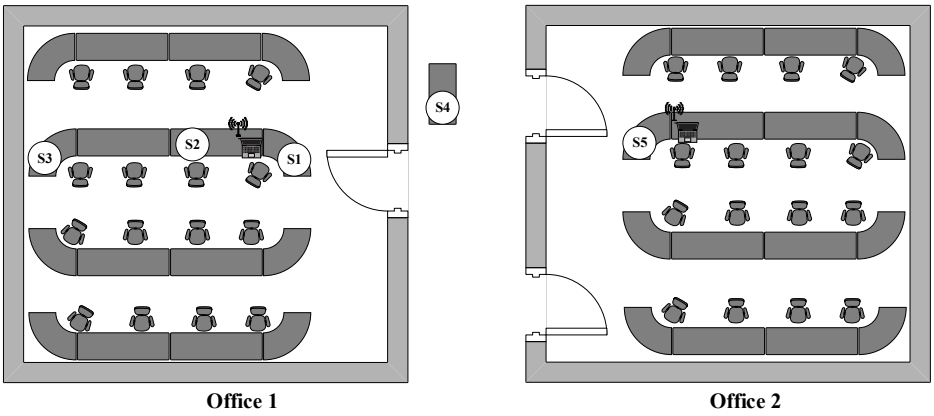

Figure 6. Illustration of indoor experimental setups.

Table 1. The dataset obtained under indoor static drone settings.

Dataset	Location	Distance (m)	Dates	Frames
S1	Office 1-site 1	0.8	9 March 2023	21,950
S2	Office 1-site 2	1	22 March 2023	18,000
S3	Office 1-site 3	4.5	7 March 2023	30,000
S4	Office 1-site 4	6	2 June 2023	4274
S5	Office 2-site 5	1	8 November 2022	24,000

There are a total of four sets of drone flight state datasets. The dataset M1 was collected while the drone was hovering at a fixed height of 10 m in the air. It should be noted that the position of the drone during hovering is not strictly fixed and may be slightly disturbed due to wind and other factors, resulting in position deviations. As shown in Figure 7, to test the performance of RFF in vertical or horizontal motion scenarios of drones, an outdoor square area M2 was used. The dataset M2-1 is collected from the drone taking off from a stationary state and staying at a fixed height of 10 m for 10 s before landing, during which the vertical climbing and landing speed is about 1 m/s. The dataset M2-2 was collected by drone in an irregular flight state at a height of 10 m in the area, with a flight speed of approximately 2~4 m/s. During this period, the remote control was used to control the drone flight in the area. Unlike the data link frequency, the control link frequency is set as 2.4 GHz, which will not cause interference to RFF. The dataset M2-3 uses ground stations for flight route planning, allowing drones to fly in circles within the area at a speed of approximately 8~12m/s. The number of signal frame samples for each device in datasets M1 and M2-2 is equal, and the number of samples for each device in datasets M2-1 and M2-3 is approximately equal. The specific information regarding the number of data frames, collection date, and transceiver distance is shown in Table 2.

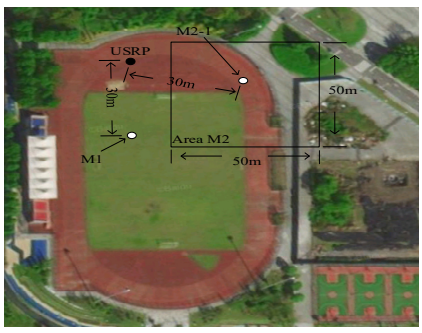

Figure 7. Illustration of outdoor experimental setups.

Table 2. The dataset obtained under outdoor hovering and flying drone settings.

Dataset	Working States	Location	Distance (m)	Date	Frames
M1	10 m hovering	M1	14	8 March 2023	3000
M2-1	10 m & landing (1 m/s)	M2	30	6 May 2023	1875
M2-2	10 m & flying (2~4 m/s)	M2	-	9 May 2023	3000
M2-3	10 m & flying (8~12 m/s)	M2	-	11 May 2023	763

The above testing scenarios involve multiple drone flight states. The drone hovering state test scenario can provide reference for studying the adverse effects caused by the fluctuation frequency $f_v(t)$ caused by drone body vibrations. The reliability of RFF in the presence of both $f_v(t)$ and $f_d(t)$ can be studied in the scenarios of vertical and horizontal motions of drones. In addition, to test the robustness of the proposed RFF method for drones in time-varying air ground channels, in the horizontal motion scenario, low speed (2~4 m/s) and high speed (8~12 m/s) were set up. With the above testing scenarios, rigorous reliability verification can be conducted on the data augmentation method and fractional frequency offset processing method.

The structure of the neural network model designed is shown in Figure 8. The model is implemented based on residual blocks and channel attention mechanisms, and uses LayerNorm regularization in the model. The input of this model is a 3×800 I-branch, Q-branch and phase sequence. The input signal data is processed through a feature extraction layer, and the output feature vector is input into a fully connected layer for classification.

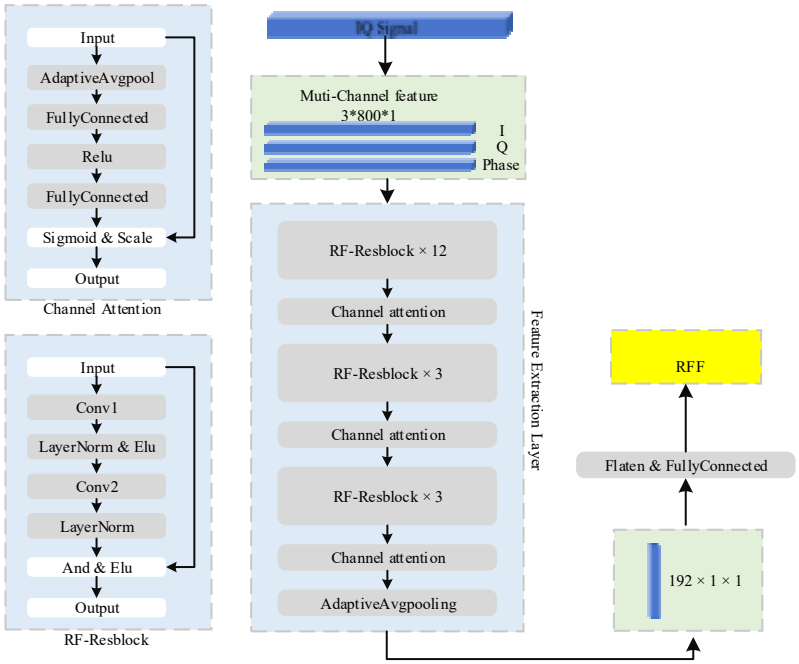

Figure 8. The neural network model used for RFF extraction.

The model training phase during the experiment was completed on a laptop hardware platform equipped with RTX 2060 GPU (Nvidia HongKong Holding. Ltd., Hong Kong, China) and Intel Core i7 9750H CPU (Intel Corporation, Santa Clara, CA, USA). At the same time, the training uses an Adam optimizer with an initial learning rate of 0.001, and the loss function uses a cross entropy function. In order to better assist the model in converging to the optimal solution, a weight decreasing strategy with a value of 0.0001 was adopted during the training process, and the training set was divided into batches of every 64 signal frames. After completing training, we use an early stop strategy. When the loss value remains unchanged for 20 consecutive epochs, terminate the training process and save the model to prevent overfitting.

4.2. Experimental Results

In the experiment, the indoor stationary state dataset S1 was used as the benchmark training set for the model. By adding Gaussian white noise, the training set S1 is subjected to noise enhancement. It should be noted that if it involves generative channel enhancement or fractional frequency offset processing, noise enhancement is performed in the final stage to prevent noise interference with the proposed method. Noise enhancement uses a random parameter enhancement method with SNR uniformly distributed, and U (10, 50) is set in the experiment. Finally, the models used are all based on the benchmark training set S1, and each method will process the benchmark training set S1 accordingly.

4.2.1. Drone Training Dataset Augmentation

In this section, for the sake of convenience, only noise enhancement is referred to as method 1, the combination of noise enhancement and fractional frequency offset processing is referred to as method 2, and the combination of generative channel enhancement, fractional frequency offset processing, and noise enhancement is referred to as the RFF method. In the experiment, the probability of channel switching 1 in the generative channel data augmentation module is set to 0.2, and the probability of switching 2 is set to 0.1. Each sample has a probability of 0.2 to skip the channel enhancement channel, and if it enters the enhancement channel, there is a probability of 0.1 to use the Rayleigh model. Considering the computational complexity, the number of sine waves in the experimental sine wave superposition method is set as $N = 50$.

To compare the proposed K factor generation strategy, a K factor generation method based on uniform distribution was selected for comparative experiments with the proposed environmental measurement method, i.e., $K \sim \mathrm{U}(min, max)$, where min and max are the maximum and minimum K values in the deployment environments, respectively. This method has a low implementation cost and only requires estimation of the maximum and minimum K factors of the deployment environment.

Figure 9 compares two sets of random parameter generation methods, including the uniform distribution method and the environmental measurement method. The figure shows that the generative channel enhancement using uniformly distributed K values results in slightly weaker average classification performance of the model. In deep learning, it is required that the distribution pattern of the training set samples be consistent with that of the test set [10]. This principle ensures that the data samples encountered by the model during training and testing have the same statistical characteristics, allowing the model to leverage the experience learned during the training phase. First, using uniformly distributed K values in the generative channel enhancement method will ultimately result in the dataset containing a large number of rare samples from real-world environments. At the same time, there are many severe fading signal samples in the dataset, which increases the complexity of the dataset and makes it difficult for the model to converge [11]. On the contrary, in data augmentation, K values generated based on environmental measurement methods are used to make the channel distribution of the training set more in line with the actual distribution in the environment, i.e., the training set and the test set have better consistency in distribution, further enabling the model to demonstrate better classification performance. Therefore, in this experiment, a K value generation strategy based on environmental measurement methods was adopted.

As shown in Figure 10, the results on drone hovering state test set M1 indicate that there is no significant difference in performance among the three methods. This is mainly because the hovering state of the drone is in a static state, which introduces body vibration compared to the true static state. But due to the speed of the drone is close to zero, it will not be affected by the fast fading effect caused by the Doppler effect. Therefore, the use of data augmentation methods on M1 has limited improvements. On the drone horizontal motion test sets M2-2 and M2-3, the fast fading channel generated by drone motion becomes the main problem of RFF, and relying solely on noise enhancement or fractional frequency offset compensation cannot solve the problem in this scenario. Thus,

the intervention of the generative channel enhancement method effectively alleviates the performance degradation problem of the model in fast fading scenarios. Compared to Method 1, the RFF method improves by 16.5% and 33% on the test sets M2-2 and M2-3, respectively. Similarly, in the takeoff and landing test set M2-1 for vertical drone movement, the proposed RFF method improves the performance by 18.26% compared to method 1. Therefore, the proposed data augmentation method based on the fast fading channel of drone can make the pre-trained neural network model more robust to the time-varying channel effects caused by the movement of drones.

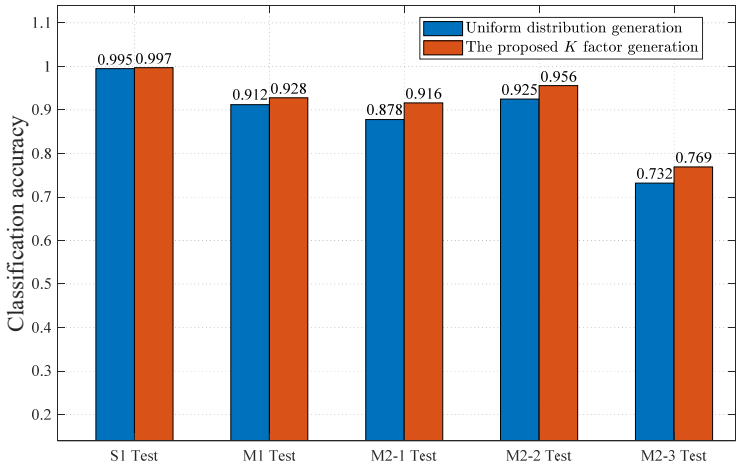

Figure 9. Classification accuracies of different parameter generation strategies.

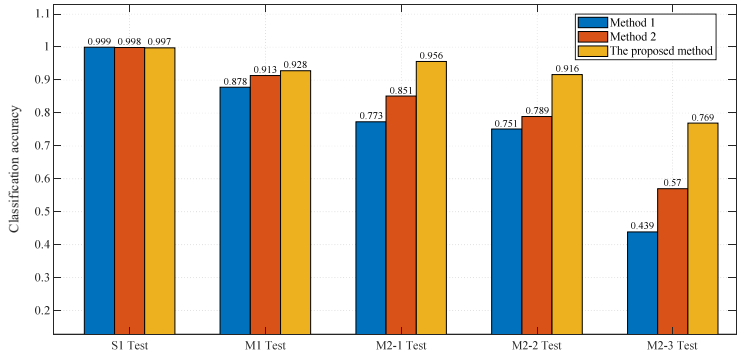

Figure 10. Performance comparisons of different methods for drone flight states.

To demonstrate the improvements of the model's generalization ability by the data augmentation module, Figure 11 compares the RFF extraction capabilities of the feature extraction layers of model Mod1 based on Method 1 and model Mod2 based on the RFF method for drone, and visualize their ability to extract RFF in dataset M2-1. The sample length of the drone preamble sequence is 800 sampling points. The feature extraction layer of this model takes a 3×800 matrix composed of I-branch, Q-branch, and phase information as inputs, and finally outputs a 192 dimensional feature vector. Due to the high dimensionality of the feature vectors, which is not conducive to visualization operations, principal component analysis (PCA) is used to reduce the dimensionality of the feature vectors to three dimensions. Obviously, model Mod1 based on method 1 in Figure 11a is no longer able to distinguish between drone data transmissions h1, h2, h3, and v2 in the scenario of drone air ground channels. The RFF of these four devices are highly confused in the feature space. However, Mod2 based on the RFF method for drone, as shown in Figure 11b, is still able to perform well in this case and can distinguish RFF of all devices, achieving a classification accuracy of 95.6%.

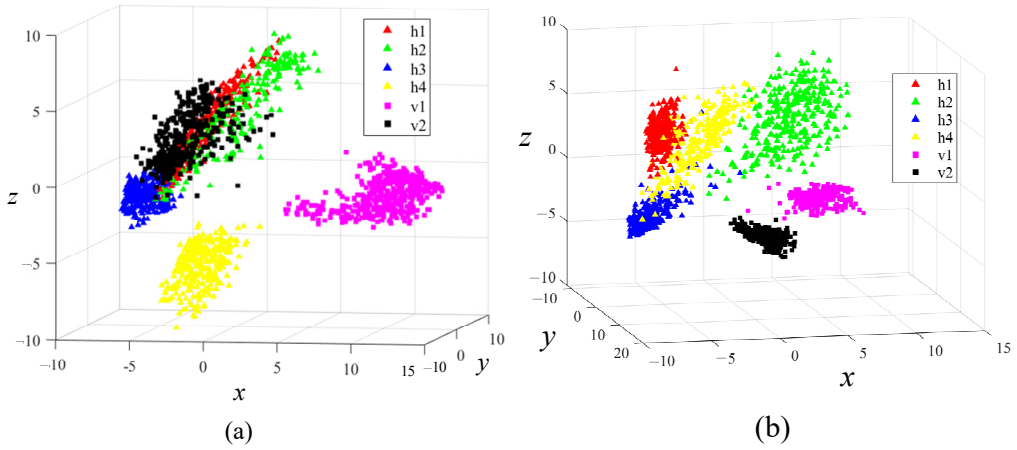

Figure 11. The fingerprint distribution of each device in the feature space of (**a**) Mod1 with classification accuracy of 77.3%, and (**b**) Mod2 with classification accuracy of 95.6%.

For test set M2-3, due to the fast drone movement speed (8–2 m/s) in the horizontal flight scenario, the air ground channel effect is the most severe among all test scenarios. Visualize the relationship between the recognition accuracy of device h2 in the test set and the *K* factor. Figure 12 shows the test results of Method 2 and the proposed RFF method, respectively. "True" is the correct classification, and "False" is the incorrect classification. Figure 12a shows that when $K < 20$ dB, Method 2 has a high model misjudgment rate and cannot reliably extract RFF. On the other hand in Figure 12b, the generative channel enhancement in the RFF method enables the model to have reliable RFF extraction capability in low *K* value scenarios. This is mainly because the generative channel enhancement step effectively improves the relatively single channel environment in the original training set. By introducing low *K* value Rician channels and a small portion of Rayleigh channels, the model exhibits high generalization ability in the actual drone low *K* value air ground channel scenarios.

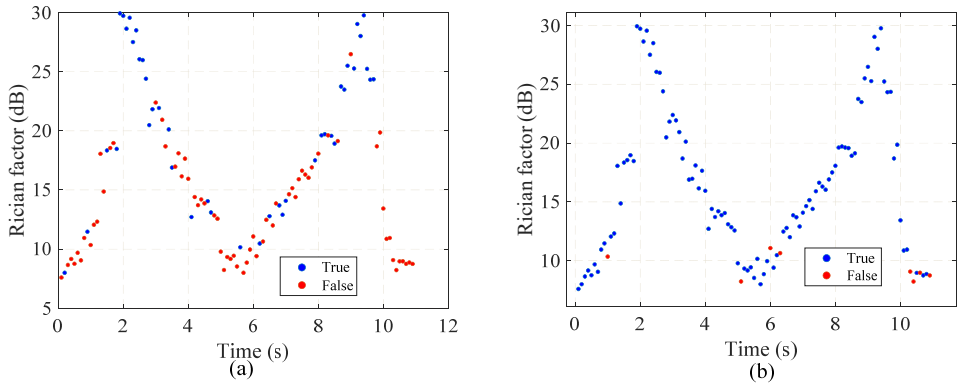

Figure 12. The *K* factor identification for (**a**) Method 2 and (**b**) the proposed method.

In addition, the convergence time of the neural network models on the training dataset processed by each method is shown in Table 3. Due to the fact that data augmentation methods essentially increase the diversity and complexity of data samples in the training dataset, the model requires more training time to achieve convergence and learn RFF features derived from the transmitter hardware from these complex patterns. In the experimental environment, a channel unbiased training set for RFF is created, and the drone can be left indoors for two minutes to complete. The use of the generative channel enhancement method avoids the endurance limitation problem and the channel distribution problem of the dataset caused by directly collecting samples from drone flight states.

Table 3. The convergence time of the training model under different methods.

Methods	Training Time
Method 1	5 min 20 s
Method 2	9 min 13 s
Proposed method	22 min 47 s

4.2.2. Drone Fractional Frequency Offset Compensation

In order to test the effectiveness of the proposed fractional frequency offset compensation method, separate tests were conducted on datasets in various static indoor environments. The static indoor scenario is considered due to evaluating purpose. Due to the absence of fast fading effects caused by drone mobility in indoor stationary environments, it can be considered that the environmental channel has a weak impact, and there are no frequency fluctuations caused by Doppler frequency shift, drone body vibration, motor rotation, and other factors. The typical relative frequency offset numerical values for Doppler shift, motor rotation, structural vibrations, temperature are roughly several hundred Hz to several thousand Hz. Therefore, there is no need to perform data augmentation on the training set S1. In a stationary state, the main influencing factors of RFF recognition are the frequency deviation changes of the crystal oscillator caused by environmental temperature or device startup time. To this end, we compare two common traditional approaches: the traditional method of preserving the complete frequency offset without processing and the traditional method of removing the complete frequency offset features. Note that the frequency offset can be estimated including the integer part and the fractional part. Without any frequency compensation, the signals are referred to as preserving the complete frequency offset, while the case that removing the complete frequency offset features refers to removing both the integer part and the fractional part. The experimental results are shown in Figure 13.

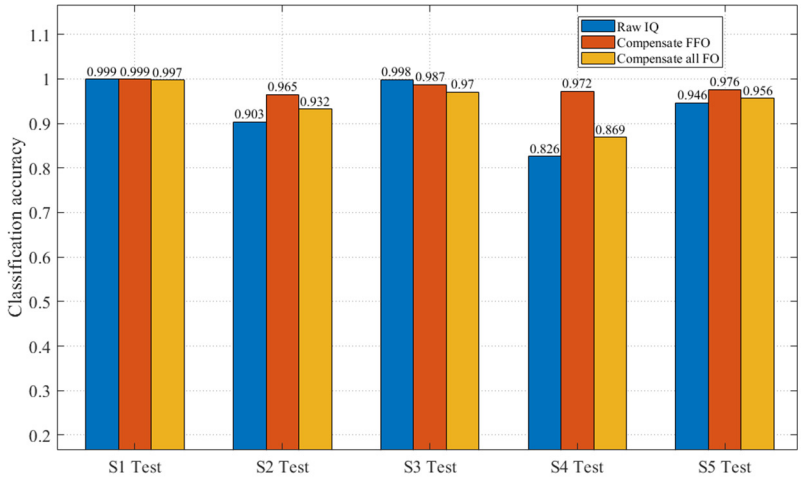

Figure 13. Comparisons of classification performance for different methods under stationary state.

It can be seen that the traditional methods occasionally achieve the best performance in the testing environment (dataset S3), but overall they exhibit an unstable phenomenon and have poorer average performance compared to other methods that deal with frequency offset. As shown in Figure 14a, the frequency offsets of all devices between dataset S1 and S3 match well, and the frequency offset features can be considered as usable RFF features between training set S1 and testing set S3. Then, the frequency offset features can be directly used for classification, the traditional Method 1 theoretically achieves better performance. As shown in Figure 14b, there is a significant difference in the frequency offset features between datasets S1 and S4 after power on. The frequency offset feature components contained in the feature space extracted by traditional Method 1 are considered

unstable and do not possess the property of becoming RFF features' short-term invariance. Although the traditional Method 2 avoids the negative impact of frequency offset caused by the environment, it also has a classification performance bottleneck due to the lack of frequency offset features. The proposed method only retains stable integer multiples in frequency offset to compensate for unstable fractional multiples caused by environmental factors, leaving behind integer multiples that often symbolize differences between models. Ultimately, by effectively utilizing frequency offset features, the classification performance surpasses the upper limit of traditional methods, while the method has stable and reliable average classification performance.

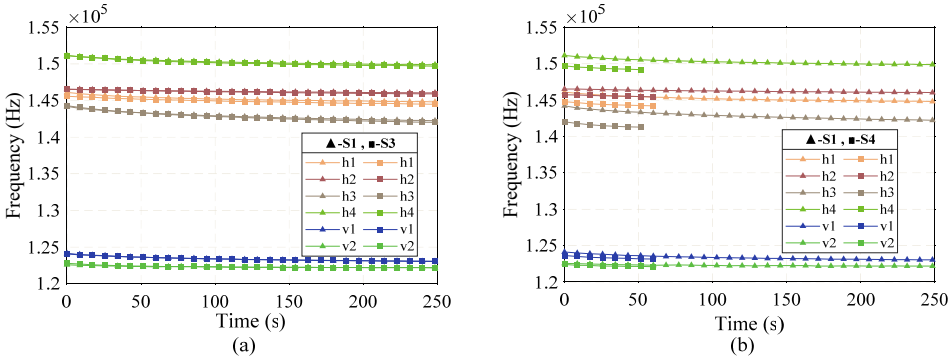

Figure 14. The frequency offsets for (**a**) dataset S1 and dataset S3, (**b**) dataset S1 and dataset S4.

To demonstrate the effective utilization of frequency offset features by the proposed fractional frequency offset processing method, Figure 15 compares the RFF extraction capabilities of the Mod3 model based on traditional Method 1 and the Mod4 model based on the proposed fractional frequency offset processing method, and visualize their ability to extract RFF in dataset S4. In Figure 15a, it can be observed that the RFF extracted by devices h1, h2, and h3 have been confused in the feature space. This is because the frequency offset characteristics of S4 devices are significantly different from those of S1 devices, which in turn affects the RFF containing the frequency offset characteristics of drones, causing drift or even overlap of the RFF of each device in the feature space. In Figure 15b, Mod4 with fractional frequency offset is suppressed, as it is no longer affected by the interference of unstable frequency offset factors on RFF, ultimately achieving a classification accuracy of 97.2%.

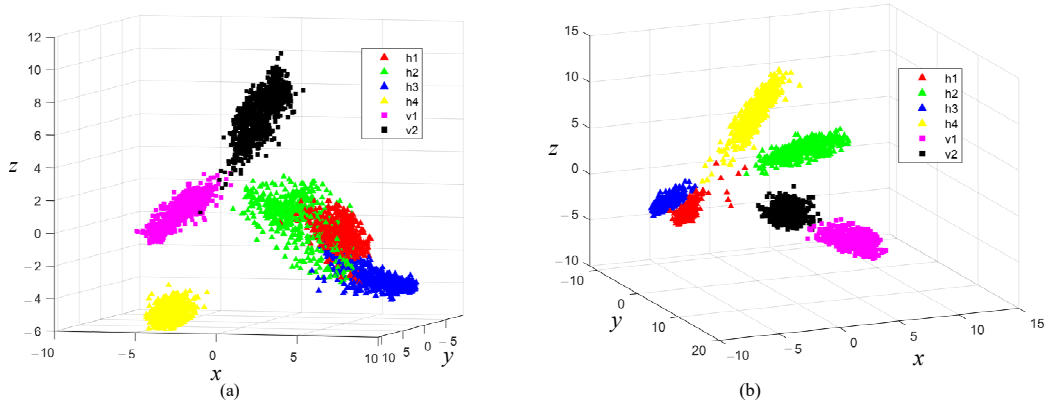

Figure 15. RFF distribution of each device, (**a**) Mod3 with classification accuracy 82.6%, (**b**) Mod4 with classification accuracy of 97.2%.

5. Conclusions

We proposed dataset augmentation and fractional frequency offset compensation based radio frequency fingerprint identification for drone communications in this work.

The training RFF dataset was augmented by regenerating the drone channel characteristics and compensate the fractional frequency offset. The proposed method estimates the Rician K value of the channel and curve-fits the statistical distribution, the Rician channels were regenerated using the sinusoidal superposition method. Experimental results show that the proposed method achieves an average classification accuracy of 97% under spatial and temporal varying conditions. Future research directions include optimization for complex attack scenarios, real-time improvement of algorithms, and integration with other security mechanisms.

Author Contributions: Conceptualization, D.L.; methodology, D.L. and Z.W.; software, Z.W. and Y.L.; validation, Z.W. and Y.L.; formal analysis, Z.W.; investigation, D.L.; resources, D.L.; data curation, Z.W. and Y.L.; writing—original draft preparation, D.L. and Y.L.; writing—review and editing, D.L. and Y.L.; visualization, D.L. and Y.L.; supervision, D.L. and H.S.; project administration, D.L.; funding acquisition, D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 62071250, the Fundamental Research Funds for the Central Universities, grant number 2242022R40030 and National Key Research and Development Project, grant number 2022YFB3904404. The APC was funded by 2242022R40030.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: Huafei Shen has no potential interests with Sumec Group Corporation, the other authors declare no conflicts of interest.

References

1. Eian, I.C.; Lim, K.Y.; Yeap, M.X.L.; Yeo, H.Q.; Fatima, Z. Wireless networks: Active and passive attack vulnerabilities and privacy challenges. *Preprints* **2020**, *21*, 230–233.
2. Hall, J.; Barbeau, M.; Kranakis, E. Radio frequency fingerprinting for intrusion detection in wireless networks. *IEEE Trans. Defendable Secur. Comput.* **2005**, *12*, 1–35.
3. Yang, X.; Li, D.M. LED-RFF: LTE DMRS-based channel robust radio frequency fingerprint identification scheme. *IEEE Trans. Inf. Forensics Sec.* **2024**, *19*, 1855–1869. [CrossRef]
4. Chalmers, C.; Fergus, P.; Curbelo Montanez, C.A.; Longmore, S.N.; Wich, S.A. Video analysis for the detection of animals using convolutional neural networks and consumer-grade drones. *J. Unmanned Veh. Syst.* **2021**, *9*, 112–127. [CrossRef]
5. Yoo, S.; Oh, H. Analysis of Commercial Drone Sounds and Its Identification. In Proceedings of the International Conference on Research in Adaptive and Convergent Systems, Gwangju, Republic of Korea, 13–16 October 2020.
6. Kunze, S.; Weinberger, A.; Poeschl, R. A Software Defined Radio Based Implementation for the Radio Frequency Analysis of Signals from Unmanned Aerial Systems. In Proceedings of the 2019 29th International Conference Radioelektronika, Pardubice, Czech Republic, 16–18 April 2019.
7. Li, D.M.; Yang, X.; Hu, A.Q.; Zhou, F.; Dobre, O.A. LTE device radio frequency fingerprints blind extraction based on temporal-frequency domain PRACH signals. *IEEE Trans. Veh. Technol.* **2023**, *72*, 13229–13242. [CrossRef]
8. Kosuke, T.; Shun, K.; Phuc, V.T.; Shinya, S.; Ahn, C.-J. Joint SNR and Rician K-factor estimation using multimodal network over mobile fading channels. *IEEE Trans. Mach. Learn. Commun. Netw.* **2024**, *2*, 766–779.
9. Lashkarian, N.; Kiaei, S. Class of cyclic-based estimators for frequency-offset estimation of OFDM systems. *IEEE Trans. Commun.* **2000**, *48*, 2139–2149. [CrossRef]
10. Hou, R.; Lo, J.Y.; Marks, J.R.; Hwang, E.S.; Grimm, L.J. Classification performance bias between training and test sets in a limited mammography dataset. *PLoS ONE* **2024**, *19*, 282402–282410. [CrossRef] [PubMed]
11. You, D.; Xiao, J.; Wang, Y.; Yan, H.; Wu, D.; Chen, Z.; Shen, L.; Wu, X. Online learning from incomplete and imbalanced data streams. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 10650–10665. [CrossRef]