


Article

A Lightweight, Secure Authentication Model for the Smart Agricultural Internet of Things

Fei Pan ^{1,2,†}, Boda Zhang ^{1,3,†}, Xiaoyu Zhao ^{1,3}, Luyu Shuai ^{1,3}, Peng Chen ^{1,3} and Xuliang Duan ^{1,2,*} 

¹ College of Information Engineering, Sichuan Agricultural University, Ya'an 625000, China; zhangboda@stu.sicau.edu.cn (B.Z.)

² Ya'an Digital Agricultural Engineering Technology Research Center, Ya'an 625000, China

³ Agricultural Information Engineering Higher Institution Key Laboratory of Sichuan Province, Ya'an 625000, China

* Correspondence: duanxuliang@sicau.edu.cn

† These authors contributed equally to this work.

Abstract: The advancement of smart agriculture, with information technology serving as a pivotal enabling factor, plays a crucial role in achieving food security, optimizing production efficiency, and preserving the environment. Simultaneously, wireless communication technology holds a critical function within the context of applying the Internet of Things in agriculture. In this research endeavor, we present an algorithm for lightweight channel authentication based on frequency-domain feature extraction. This algorithm aims to distinguish between authentic transmitters and unauthorized ones in the wireless communication context of a representative agricultural setting. To accomplish this, we compiled a dataset comprising legitimate and illegitimate communication channels observed in both indoor and outdoor scenarios, which are typical in the context of smart agriculture. Leveraging its exceptional perceptual capabilities and advantages in parallel computing, the Transformer has injected fresh vitality into the realm of signal processing. Consequently, we opted for the lightweight MobileViT as our foundational model and designed a frequency-domain feature extraction module to augment MobileViT's capabilities in signal processing. During the validation phase, we conducted a side-by-side comparison with currently outstanding ViT models in terms of convergence speed, precision, and performance parameters. Our model emerged as the frontrunner across all aspects, with FDFF-MobileViT achieving precision, recall, and F-score rates of 96.6%, 95.6%, and 96.1%, respectively. Additionally, the model maintains a compact size of 4.04 MB. Through comprehensive experiments, our proposed method was rigorously verified as a lighter, more efficient, and more accurate solution.

Keywords: smart agriculture; information security; wireless communication technology; Internet of Things



Citation: Pan, F.; Zhang, B.; Zhao, X.; Shuai, L.; Chen, P.; Duan, X. A Lightweight, Secure Authentication Model for the Smart Agricultural Internet of Things. *Agronomy* **2023**, *13*, 2257. <https://doi.org/10.3390/agronomy13092257>

Academic Editor: Magdalena Sobocińska

Received: 26 July 2023

Revised: 24 August 2023

Accepted: 25 August 2023

Published: 28 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In smart agriculture, wireless communication technology, as the cornerstone of the agricultural IoT, will be used in much automation equipment. Unlike wired communication, wireless communication is not bound by wires. There is no need to consider issues such as wiring, distance, link maintenance, troubleshooting, etc., and the low maintenance cost makes it more suitable to deploy in agricultural automation equipment. Therefore, wireless communication is crucial for developing smart agriculture [1–4].

Data privacy and security are the intellectual core of the Internet of Things (IoT) [5], and it is important to study how secure communication in the IoT can be ensured. To resist eavesdropping and attacks by illegal users and to guarantee the security and reliability of communications, most communication systems use traditional cryptography-based techniques to ensure the security of communications, such as the PKI-based security protocol for connected vehicles proposed by Raya et al. [6], where each vehicle is equipped

with multiple private keys and their corresponding certificates; however, the PKI security protocol is inefficient. It can only meet the needs of a few vehicles due to its computational difficulties. Biswas et al. [7] proposed an ID-based proxy approach using signatures, which could improve efficiency, but the keys are prone to leakage. Asim et al. [8] proposed a hierarchical attribute-based cryptographic system (HABE) in which the key acts as a leaf, but if the root authority is leaked, the whole system is corrupted. Moreover, the implementation of the HABE leads to high computational overhead. In the current smart agriculture IoT, most wireless automated sensors and smart devices must maintain low-power operation when unmonitored. The algorithmic complexity of traditional cryptography for key distribution and management is high, which results in high energy consumption and reduced lifetimes for nodes [9], so we need authentication methods that are more suitable for smart agriculture.

Currently, more and more researchers are tending to use channel characteristics for secure authentication. According to electromagnetic propagation theory and the Jakes spectral channel model, the channel fading between each transmitter and receiver is independent when the distance between different transmitters is better than half a wavelength [10], and the channel between the same pair of transceivers changes slowly during its coherence time, when the channel is both spacetime-unique and challenging to forge. For example, Shabnam Shafiee et al. [11] employed the multiple-input single-output (MISO) technique to use channel state information in different antennas with the same information for beam assignment, which was kept as close as possible to the main channel direction. Dania Marabissi et al. [12] studied PHY layer continuous authentication and spoof detection and proposed a machine learning wireless fingerprint identification method for wireless sensor networks (WSNs) that exploits the features of the channel information to verify whether the current message is reliable or not.

Traditional authentication methods usually employ signal processing techniques, such as computational-spectrum, time-domain, and frequency-domain analyses, to detect the possibility of illegal signals. These methods rely mainly on pre-written rules and thresholds to identify illegal signals and thus cannot learn and adapt to new illegal signal types and can be limited by environmental noise and a lack of signal diversity. In contrast, deep learning algorithms can adapt to different data types and tasks [13–16]. Yadava et al. [17] summarized the application of federated learning in privacy preservation and discussed in detail the encryption mechanisms based on federated learning. As a subset of deep learning, the utilization of federated learning in privacy preservation indirectly demonstrates the viability of training deep neural networks for channel authentication through the utilization of channel feature data. Moreover, deep learning algorithms can learn using various smart sensors running on massively distributed massive datasets, and they show better adaptive and generalization capabilities under different environmental conditions. This study applies deep learning techniques to the channel authentication task in smart agriculture settings with complex environments and different device functions and locations to obtain better authentication accuracy.

As Transformers [18] have made a big splash in the field of computer vision, more and more researchers are utilizing Transformers as a network backbone or drawing on their ideas to improve their algorithms. Transformers are also well suited for use in signal processing. Transformers leverage self-attention mechanisms to capture long-range dependencies within signals, whether temporal sequences or spatial patterns. This enables them to better capture correlations between different parts, ultimately enhancing processing accuracy. Additionally, the Transformer adapts well to signal patterns of varying scales and frequencies, showcasing exceptional performance when handling multiscale signals. During its training process, the Transformer automatically learns representations suited to the unique characteristics of signals. It can directly extract crucial information from raw signals, significantly streamlining the signal-processing workflow and leading to a more automated and efficient process. Furthermore, the Transformer boasts the advantage of parallel computing, enabling it to handle large-scale signal data efficiently. This

characteristic proves highly beneficial for real-time signal processing and managing extensive signal datasets, effectively accelerating the entire processing pipeline and enhancing system responsiveness. Therefore, this study used MobileViT [19] as the backbone. MobileViT combines convolution with the Transformer to construct a lightweight backbone network. This study proposes a new channel authentication model called FDFE-MobileViT for agricultural environments based on MobileViT.

The main contributions of this study are:

- (1) A frequency-domain feature extraction (FDFE) module utilizing frequency-domain information with a signal plus windowing is proposed;
- (2) A new channel authentication model called the FDFE-MobileViT network for smart agriculture environments is proposed;
- (3) The authentication effect of FDFE-MobileViT was experimentally verified in smart agriculture indoor and outdoor environments.

2. Materials and Methods

2.1. Agricultural IoT Model

As shown in Figure 1, this study considered a smart agriculture scenario including intelligent terminals, drones, robotic arms, etc. As in the indoor environment shown in Figure 1a, the drone was equipped with infrared sensors, multispectral sensors, etc. The data acquired by the drone were uploaded as high-definition images and sensor data. The robotic arm could link with the drone data according to the crop's needs and automatically complete planting and pruning. Smart sensors all over the indoor environment providing real-time monitoring of light intensity, air humidity, and carbon dioxide concentration transmitted the data to the central control system. In the outdoor environment shown in Figure 1b, the plant protection drone, innovative farming machine, and intelligent irrigation system undertook pesticide spraying, field finding, crop harvesting, and irrigation in the farmland according to the central control system. Since different devices have different movement states and may be stationary or in motion, they usually use wireless networks to transmit information conveniently.



Figure 1. Smart agriculture communication scenario. (a) Indoor Environment-Plant Factory (b) Outdoor Environment-Smart Farming.

This study simplifies the situation to the communication model shown in Figure 2, where a receiver Alice communicates with multiple transmitters Bob1, Bob2, . . . , Bobn, and a masquerader, Mallory, tries to imitate Bob to send data to Alice to control Alice. Alice, Bob, and Mallory are placed at different locations when Alice needs to verify the received data. Assuming an orthogonal frequency division multiplexing (OFDM) system is used, the relationship between transmitted and received data is given by Equation (1):

$$Y = HX + b \quad (1)$$

where Y represents the received data, X represents the sent data, b represents the noise data, and H is the channel state information. Compared with Mallory, the relative positions of Alice and Bob are fixed in the stationary state, and their trajectories are relatively regular in the moving state. Therefore, the channel state information between Alice and Bob has a relatively fixed structure. At the same time, the position of Mallory is challenging to determine, so the use of channel state information can enable Alice to distinguish between Bob and Mallory.

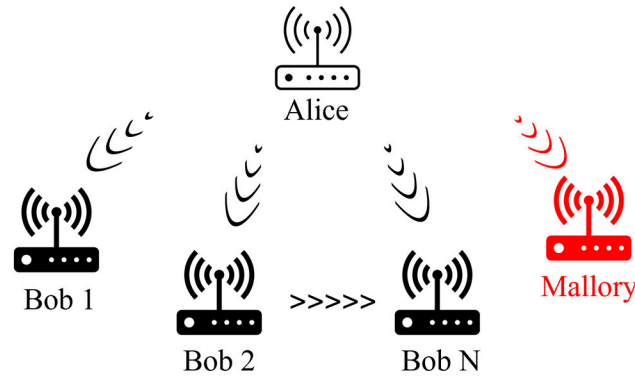


Figure 2. System model.

2.2. Data Acquisition

2.2.1. Acquisition Equipment

An NI USRP-2943R was the hardware device used in this experiment as the transmitter and receiver, and it was mainly responsible for receiving and transmitting the signals in the wireless environment and completing data conversion. The software used in this study was LabVIEW2019-SP1 with a 2×2 MIMO-OFDM system, as shown in Figure 3.

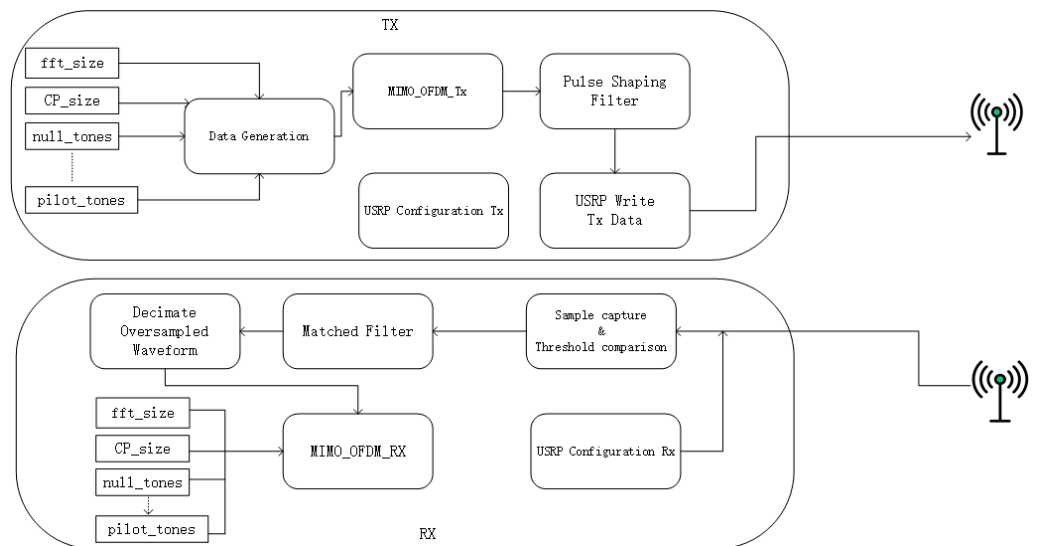


Figure 3. OFDM system structure diagram [20].

2.2.2. Acquisition Environment

Plant stands, sensors, etc., are usually installed in plant factories to monitor and control the plant growth environment to maximize plant growth [21]. However, these objects and devices can have some effects on wireless signals. Plant stands may cause obstruction and interference for wireless signals. As plant stands are usually made of metal and have a certain mass and density, they can obstruct and attenuate the wireless signals to some extent, thus reducing the strength and stability of the signals [22]. The shape and placement of the plant stand can also cause signal reflection and interference,

which further affects the signal quality. The circuitry and signal processor in the sensor may generate electromagnetic interference that interferes with the surrounding wireless signals. The sensor also uses wireless communication to transmit the collected data to the data processing equipment, which takes up some wireless channel resources and further affects the stability of the wireless signal. Plants may also affect the wireless signal; giant plants and green walls can block the signal propagation path, thus leading to some interference and attenuation of the signal [23]. In addition, the plants themselves may significantly interfere with the electromagnetic signal when the plants are thriving, and the interference with the signal may be more apparent [24].

To simulate a real-life plant environment, a smart agricultural plant factory was specially built in this study, as shown in Figure 4a, and the USRP was used to collect legal and illegal wireless channel data inside this plant factory. The receiver was placed in a corridor surrounded by plant shelves, and the transmitters were placed in four stationary positions and on two motion tracks, as shown in Figure 5a.



Figure 4. Acquisition environment: (a) indoor, (b) outdoor. (Red boxes are receivers and transmitters).

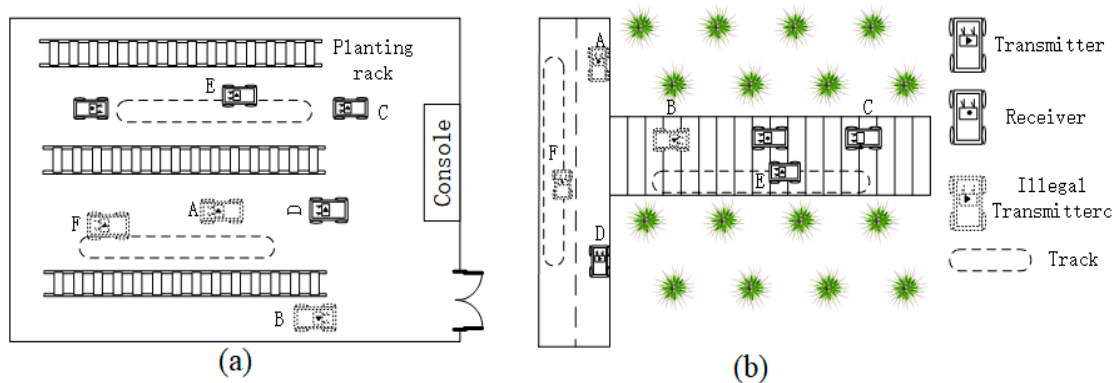


Figure 5. USRP placement in indoor and outdoor situations: (a) indoor, (b) outdoor.

An outdoor farmland environment will have some influence on the transmission of wireless signals. There are usually a lot of vegetation, soil, and water bodies in the farmland environment. Their physical characteristics interfere with and attenuate the transmission of wireless signals. Some agricultural machinery facilities are utilized in the farmland environment, such as agricultural machinery and large irrigation equipment, which generate significant electromagnetic radiation, and the movement of these facilities also leads to signal fluctuations. In addition, some radio frequency interference sources in the farmland environment, such as agricultural remote sensing devices, wireless monitoring equipment, sensors, etc., can interfere with the wireless signals and affect the transmission

and reception quality. The impact of the farmland environment on the wireless signal is mainly reflected in signal attenuation, fluctuation, and interference.

To collect the channel data from the outdoor farming environment, the transmitter and receiver devices were placed in a farming field in Ya'an, Sichuan province, China, as shown in Figure 4b. The receiver was placed at a fixed position in the field canyon and the transmitters were placed at four stationary positions and on two moving tracks, as shown in Figure 4b. Considering the complexity of the agricultural environment, this study collected outdoor data under sunny, cloudy, light rain, daytime, nighttime, and dry versus wetland conditions.

2.2.3. Data Processing

In this experiment, the channel data were collected for two cases: indoor and outdoor. During the acquisition process, the receiver was kept stationary. The transmitter was stationary or in a state of uniform motion. The transmitted data at one location were artificially determined as illegal data, and the data at the other location were legal data during the data labeling stage. According to Figure 5a,b, the data sent by the transmitter at the two positions A and B and on the F motion track were illegal, and the data sent by the transmitter at the two positions C and D and on the E motion track were legal. No detailed distinction was made between the data, which were uniformly divided into legal and illegal data. After filtering, the total number of indoor static legal channels was 2906 frames, the total number of illegal channels was 2896 frames, the total number of indoor dynamic legal channels was 2894 frames, and the total number of illegal channels was 2916 frames. The total number of outdoor static legal channels was 2830 frames, the total number of illegal channels was 2832 frames, the total number of outdoor dynamic legal channels was 1915 frames, and the total number of illegal channels was 1533.

The steps for converting the channel data to images were as follows. (1) Firstly, the training sequence part of the received signal was extracted. The fundamental part was separated from the imaginary part. (2) Most of the fundamental part of the static data was in the range from -1 to 1 . In contrast, the fundamental part of the dynamic data had a more extensive range. Therefore, the fundamental part was first normalized to map the data to $0\sim 1$ and then multiplied by 255 to make the pixel value larger for the image's red channel. The imaginary part underwent the same operation for the green channel of the image and the blue channel was all set to zero. (3) The three channels were spliced to get the channel RGB image, as shown in Figure 6.

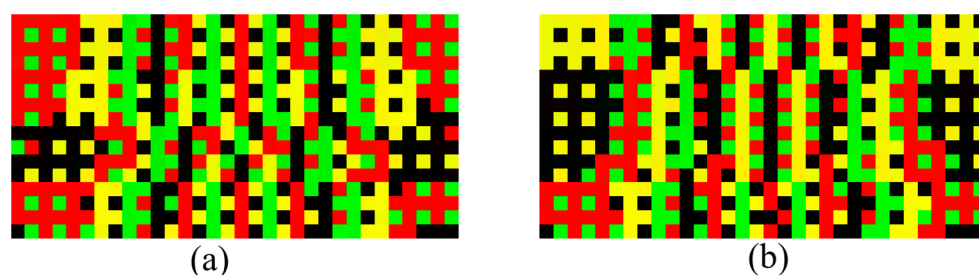


Figure 6. Transformed channel data: (a) legal channel example, (b) illegal channel example.

2.3. Model Architecture

2.3.1. Model Architecture Overview

MobileViT is a lightweight Transformer-based vision model, which is smaller and faster than other lightweight ViT models and performs better than other ViT models in various vision tasks. The accuracy of MobileViT is comparable to that of other ViT models for various vision tasks. Since MobileViT has fewer parameters, it can be more easily loaded in a device. Moreover, MobileViT is optimized for low computing power and low-power-consumption devices, and its design process optimizes the model architecture and training methods for the characteristics of a device's CPU, GPU, and other hardware platforms, thus further improving the model's operational efficiency while ensuring accuracy. As a

result, MobileViT has significant advantages over other lightweight ViT models in terms of model size, inference speed, portability, and optimization for low-arithmetic applications, which also allows it to provide a feasible solution and technical support for the channel authentication task in smart agriculture application scenarios.

This study used MobileViT as the backbone network for feature extraction of channel state information. However, since the MobileViT block only fuses features within layers, there is no shortcut branching between layers, and the frequency-domain information specific to the signal is not utilized. Therefore, this study proposed to use the frequency-domain feature extraction (FDFE) module to improve MobileViT for signal processing and obtained the FDFE-MobileViT model. Figure 7 shows the overall model architecture of FDFE-MobileViT.

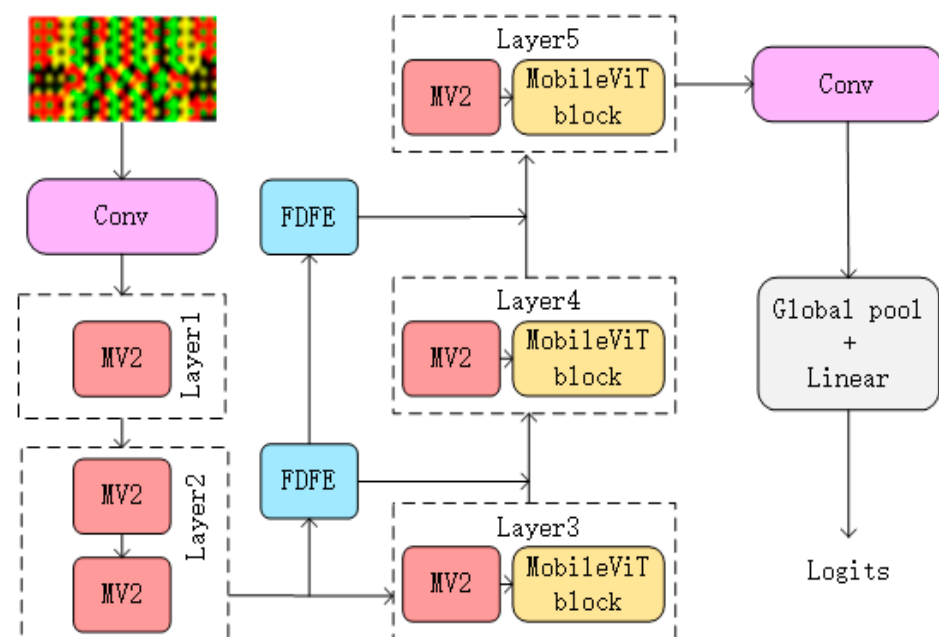


Figure 7. Model structure of FDFE-MobileViT.

Traditional attention mechanisms and feature extraction modules seldom deal with frequency-domain features [25–28]. Nevertheless, the signal needs to be analyzed in the frequency domain to obtain the spectral characteristics and frequency components, so we propose the FDFE module to supplement the shortcomings of the existing attention mechanisms. Compared with the traditional attention mechanisms, the FDFE module can filter the signal and retain the original information, so the model can simultaneously obtain the signal time-domain and frequency-domain features. The FDFE module consists of four parts: FFT, Add Windows, DSC Conv, and AvgPool. FFT transforms the image from the time domain to the frequency domain so that DSC Conv can extract frequency-domain features. The Add Window operation is similar to frequency correction and filtering, and Avg Pool downscales the features to make the output of the FDFE more compact.

This model places the FDFE module between the existing structures with shortcut-branching-layer hopping links allowing the output of Layer2 and Layer3 to be utilized by Layer4 and Layer5 so that each MobileViT block layer can get more information and obtain better results.

2.3.2. Encoder Section

The body of FDFE-MobileViT consists of Conv, MV2, FDFE, and MobileViT blocks.

The initial layer of FDFE-MobileViT is a convolutional layer with a convolutional kernel size of 3×3 . It turns the width and height of the data entering the network into one half of the original data.

The signals sent by standard communication systems can be treated as nonperiodic signals, which do not show repeating patterns over time. When the Fourier transform is used, the nonperiodic signal becomes a superposition of infinitely periodic sine and cosine functions representing discrete signals in the time domain. After the Fourier transforms, the features that are difficult to analyze in the time domain can be demonstrated in the frequency domain. Therefore, the FDFE first uses the two-dimensional discrete Fourier transform to transform the discrete signal in the time domain into the frequency domain to extract the frequency domain features. A given tensor $X_T \in \mathbb{R}^{H \times W \times C}$, X_T passes through the FFT module, which then transforms X_T into the frequency domain to obtain X_F . The transformation process is given by Equation (2):

$$X_F = \sum_{m=0}^{M-1} X_T e^{-j2\pi Fm/M}, F = 0, 1, \dots, M-1 \quad (2)$$

Here, j denotes the imaginary unit; F denotes the subscript of the frequency domain sampling point, ranging from 0 to $M-1$; and m denotes the subscript of the time domain sampling point, ranging from 0 to $M-1$.

Usually, the imaginary part generated by the Fourier transform of the image is ignored. Still, the signal naturally contains both real and imaginary parts in the initial state for non-periodic signals, so the imaginary part also carries information. Therefore, the imaginary part is not discarded in the FDFE module. Nevertheless, the real and imaginary parts are generated as separate channels in the data preprocessing; i.e., the imaginary and real parts are treated together as frequency-domain features.

To eliminate the interference of noise and make the frequency-domain data pure for subsequent extraction of frequency-domain features, FDFE uses the windowing operation. Adding a window to a signal involves multiplying the original signal by a window function to reduce the leakage of the signal in the frequency domain. A filter can achieve a similar operation, but the filter processes or changes the signal characteristics by removing or changing the signal in a certain frequency range. In contrast, during signal windowing, the signal itself is not filtered and only the amplitude is adjusted. In some cases, signal windowing can be considered a special form of filtering. The FDFE in this study used the windowing operation shown in Equations (3) and (4) for X_F :

$$W_n = 1 - \left| \frac{2n}{M-1} - 1 \right| \quad (3)$$

$$X_{w_n} = X_{F_n} * W_n \quad (4)$$

where W_n is the set of Bartlett windows; X_{F_n} and X_{w_n} are the n th elements of X_F and X_w , respectively; and X_w is the result for X_F after the windowing operation.

After the frequency-domain conversion and windowing operation, the frequency-domain features of the channel image will be more prominent, and Figure 8 shows the channel image after the operation.

Figure 8c,f show the feature maps generated by the legitimate and illegitimate channels after the window. Figure 8c,f further widen the gap between the grayscale values compared to the unwinded Figure 8b,e.

Table 1 shows the pixels' median, mean, variance, and standard deviation from the data. The significant increase in the variance and standard deviation indicate that the data were less concentrated around the mean after the windowing operation. The increase in the mean indicates that the overall data distribution was skewed toward larger values, and the increase in the median indicates that the center of the pixel distribution shifted upward, as shown in Figure 9. Thus, after the windowing operation, it is possible to make the different features of the channel images more obvious.

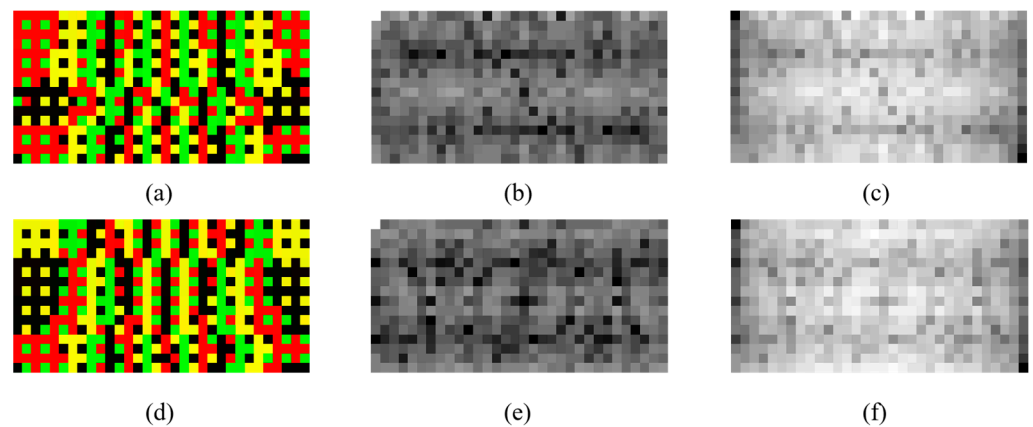


Figure 8. Comparison before and after adding windows: (a–c) results for the legal channel after frequency-domain conversion and windowing, (d–f) results for the illegal channel after frequency-domain conversion and windowing.

Table 1. Median, mean, variance, and standard deviation for the channel image pixels after adding windows (ICBAW: illegal channel before adding windows, ICAAW: illegal channel after adding windows, LCBAW: legal channel before adding windows, LCAAW: legal channel after adding windows).

| Channels | Statistics | Median | Mean | Variance | Standard Deviation |
|----------|------------|--------|-------|----------|--------------------|
| ICBAW | | 0.401 | 0.382 | 0.017 | 0.132 |
| ICAAW | | 0.764 | 0.734 | 0.023 | 0.154 |
| LCBAW | | 0.431 | 0.415 | 0.017 | 0.131 |
| LCAAW | | 0.745 | 0.719 | 0.023 | 0.152 |

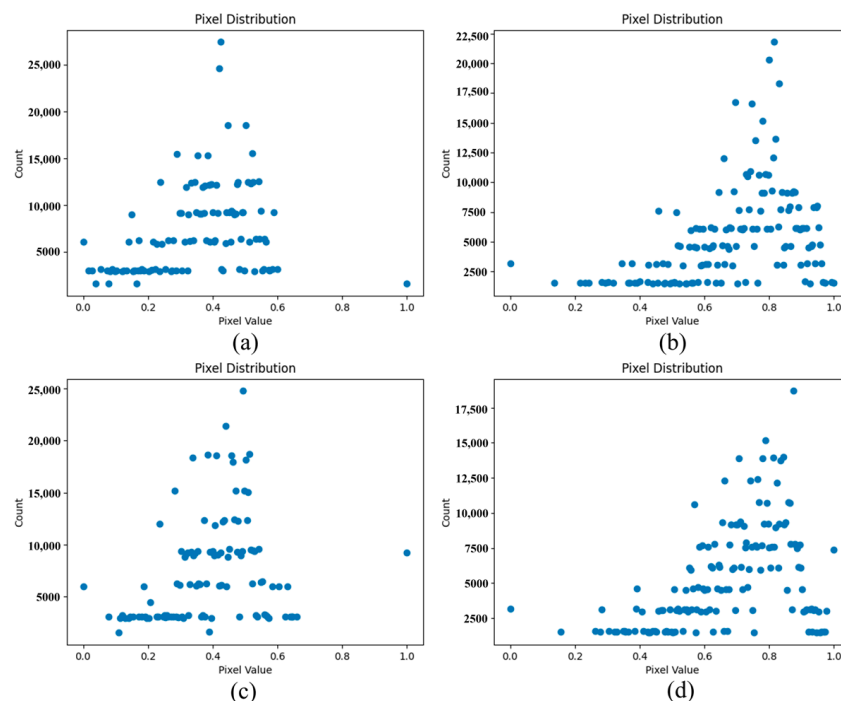


Figure 9. Changes in pixel distribution after adding windows: (a) illegal channel before adding windows, (b) illegal channel after adding windows, (c) legal channel before adding windows, (d) legal channel after adding windows.

After windowing, the frequency domain features of the image were more obvious, but the features were still too scattered, so further feature extraction was needed.

Howa A G et al. [29] proposed DSC Conv in 2017, which extracts features through two parts, Depthwise (DW) and Pointwise (PW), and has a lower number of parameters and lower operational cost compared to the conventional convolution operation. Equations (5) and (6) show the process of extracting features using DSC Conv for X_w :

$$X_c = X_{w_c} \times K_c \tag{5}$$

$$Y_F = \sum_{c=1}^{C_{in}} X_c \times P_c \tag{6}$$

where K_c indicates that a convolution kernel of size $k_w \times k_h$ is used for a tensor with a number of output channels c . X_{w_c} indicates the data of the c th channel in the input X_w . P_c is a convolution kernel of size 1×1 , which performs linear combination operations on each output channel.

Figure 10a–i show the feature maps obtained from the first feature extraction module DSC Conv, with each map corresponding to a channel, and Figure 10m shows the waveforms corresponding to each feature map. It can be seen that the waveforms corresponding to Figure 10e,j had higher amplitudes, which was reflected in the feature maps. In contrast, the waveforms corresponding to Figure 10d,f,k had lower amplitudes, which was reflected in the feature maps with lower gray values.

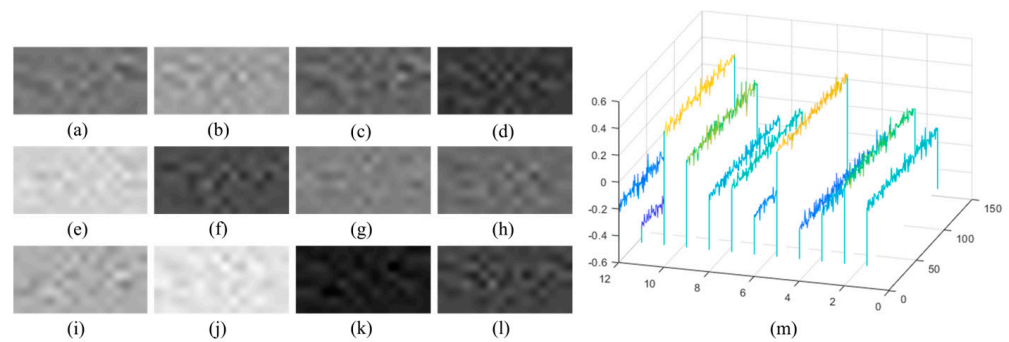


Figure 10. (a–l) Feature maps corresponding to the different channels after convolution, (m) waveforms corresponding to the different channels.

Finally, the obtained feature maps were pooled according to the average, and the specific steps are given by Equation (7):

$$Y_A = \frac{1}{k_w \times k_h} \sum_{m=0}^{M-1} Y_F \tag{7}$$

where Y_A is the output result of the FDFE module.

In FDFE-MobileViT, the FDFE feature extraction module is placed between Layer2 and Layer4 and between Layer3 and Layer5, as shown in Figure 11.

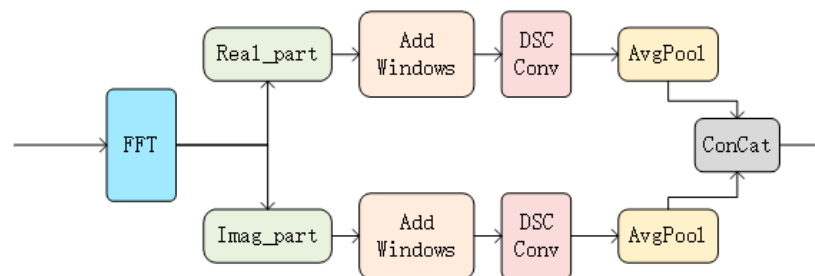


Figure 11. Structure of FDFE feature extraction module.

2.3.3. Decoder Section

The composition of the decoder was relatively simple, consisting of Conv, Global Pool, and Linear. The Conv layer utilized a convolution of size 1×1 to change the channels of the features for subsequent operations. The Global Pool layer performed dimensionality reduction on the feature map. The Linear layer was used as a classifier.

3. Experimental Process

3.1. Realization Details

In this study, the hardware used included an NVIDIA Quadro RTX 5000 graphics processing unit (GPU) (NVIDIA Santa Clara, CA, USA) with 16 GB of video memory and 128 GB of RAM, and the central processing unit (CPU) was an i9-10900K at 3.7 GHz. The system was Windows Server 2019 Standard with PyTorch version 2.0 and Python version 3.10. To verify the effectiveness of the FDFE module, this study selected the more prominent contemporary ViT models Swin-Transformer [30], NextViT [31], and T2T-ViT [32]; the lightweight model Mobilenetv3 [33]; and the benchmark ViT model ConvNexT [34] and the benchmark model MobileViT for comparison experiments using similar sized models. The loss function used a cross-entropy function. The calculation procedure is given by Equation (8):

$$Loss = -\frac{1}{N} \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (8)$$

where N is the number of samples in the dataset, y_i denotes the i th sample, and \hat{y}_i denotes the prediction probability of the i th sample.

3.2. Training Networks

FDFE-MobileViT was implemented in the Pytorch 2.0 environment and trained using the server mentioned in Section 3.1; the optimizer was an SGD optimizer with default parameters trained with 200 epochs, a batch size of 64, and an initial learning rate of 0.001, and the learning rate was adjusted using the cosine annealing method. The computation is given by Equation (9):

$$lr = lrf + \frac{1}{2}(lri - lrf) \left(1 + \cos\left(\frac{T_{cur}}{T_{max}} \pi\right) \right) \quad (9)$$

where lr denotes the current learning rate, lrf denotes the final decay rate of the learning rate, lri denotes the initial learning rate, T_{cur} denotes the current number of training rounds, and T_{max} denotes the total number of training rounds. The learning rate was changed according to the number of training rounds to achieve a better training effect.

4. Discussion

4.1. Model Accuracy Comparison

In this experiment, the collected legal and illegal data were divided into training and validation sets in a ratio of 8:2. Figures 12 and 13 show the top training and validation curves for the ConvNexT, MobileViT, NextViT, T2T-ViT, Swin-Transformer, Mobilenetv3, and FDFE-MobileViT models with the smart agriculture channel authentication dataset. It can be seen that FDFE-MobileViT had higher initial accuracy, faster convergence, and a lower oscillation amplitude than the other networks, and the accuracy was maintained at a high level during the validation process, indicating that the stability of FDFE-MobileViT was relatively high.

Table 2 shows the five networks' precision, recall, and model size with the dataset for the legal and illegal channels for smart agriculture. From Table 1, we can see that FDFE-MobileViT had 7%, 6.8%, and 7% higher precision, recall, and F-score, respectively, compared to the unimproved MobileViT with only a 0.28 MB increase in model size.

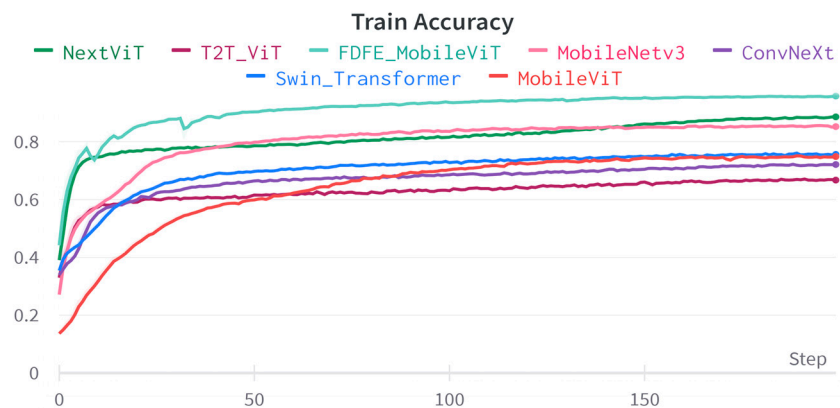


Figure 12. Training accuracy curves.

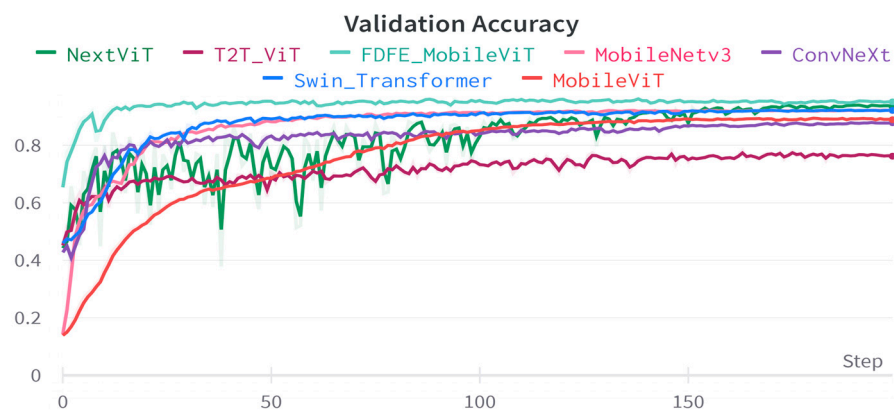


Figure 13. Validation accuracy curves.

Table 2. Precision, recall, F-score, and model size for seven networks with the dataset for the legal and illegal channels for smart agriculture.

| Models | Precision (%) | Recall (%) | F-Score (%) | Model Size (MB) |
|------------------|---------------|------------|-------------|-----------------|
| FDFE-MobileViT | 96.6 | 95.6 | 96.1 | 4.04 |
| MobileViT | 89.6 | 88.8 | 89.1 | 3.76 |
| Mobilenetv3 | 92.7 | 91.3 | 92.0 | 5.93 |
| NeXtViT | 91.8 | 90.9 | 98.9 | 117 |
| T2T-ViT | 86.9 | 86.0 | 98.1 | 15.3 |
| ConvNeXt | 88.6 | 87.9 | 88.2 | 106.2 |
| Swin-Transformer | 92.8 | 92.0 | 92.4 | 105.2 |

4.2. FDFE Module Internal Ablation Experiments

To ensure that each component of the FDFE module contributed effectively to feature extraction, this study conducted ablation experiments within the FDFE framework. These experiments individually validated the utilization of real and imaginary parts of the data, the effectiveness of the signal windowing component, and the impact of DSC Conv. The internal ablation experiments within the module were divided into three parts:

1. Utilization of only the real part obtained from the FFT in the FDFE module compared to the utilization of both real and imaginary parts;
2. Removal of the signal windowing component to contrast its effects with the process with windowing included;
3. Substitution of DSC Conv with a regular 2D convolution to compare the performance and parameter count.

Replacing the DSC Conv within the FDFE module with a standard 2D convolution, as depicted in Figures 14 and 15, yielded comparable results. Furthermore, as shown in

Figures 16 and 17, the convergence rates of both methods also tended to align. However, as indicated in Table 3, adopting the DSC Conv led to a reduction in model parameters. Hence, the use of DSC Conv holds significance.

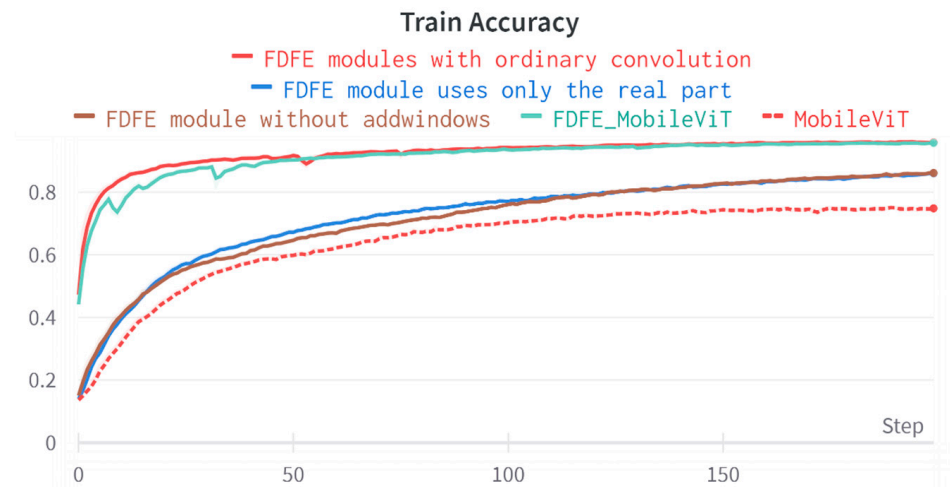


Figure 14. Module ablation experiment: training set accuracy curve.

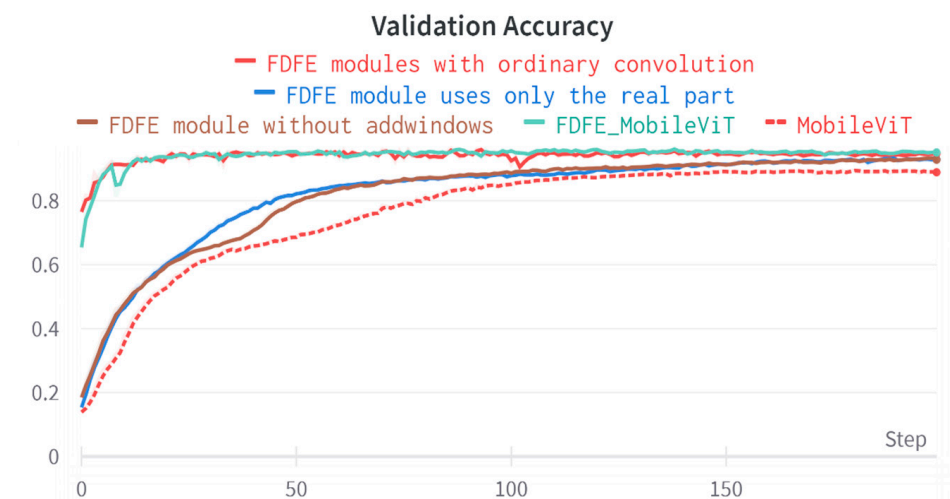


Figure 15. Module ablation experiment: validation set accuracy curve.

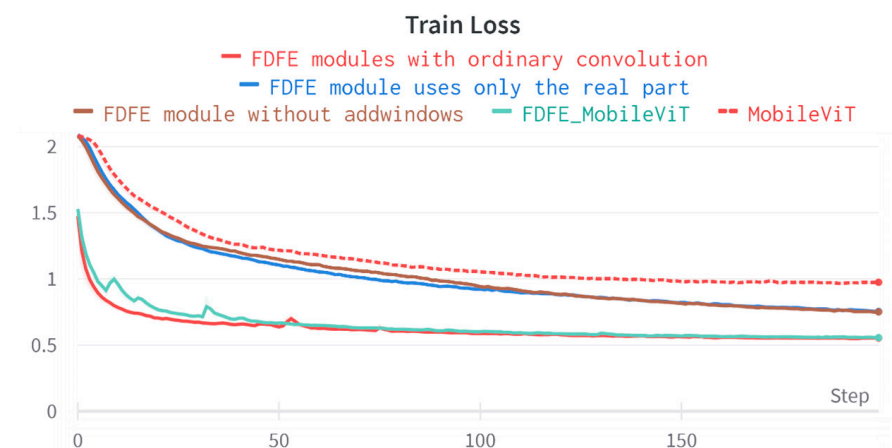


Figure 16. Module ablation experiment: training set loss curves.

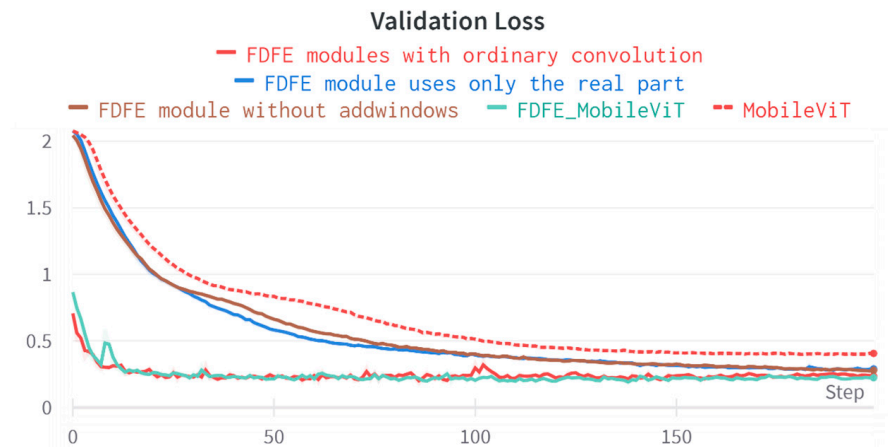


Figure 17. Module ablation experiment: validation set loss curves.

Table 3. Precision, recall, F-score, and parameters of different FDFE modules.

| Models | Precision (%) | Recall (%) | F-Score (%) | Params (MB) |
|-----------------------------|---------------|------------|-------------|-------------|
| FDFE-MobileViT | 96.6 | 95.6 | 96.1 | 1.02 |
| Ordinary convolution | 96.7 | 95.2 | 95.9 | 1.04 |
| Only the real part | 93.9 | 93.2 | 93.5 | 1.02 |
| Without addition of windows | 93.7 | 93.3 | 93.4 | 1.02 |
| MobileViT | 89.6 | 88.8 | 89.1 | 0.95 |

In this experiment, utilizing only the real component and omitting the windowing operation failed to achieve the effectiveness of combining real and imaginary components along with windowing. The conclusions drawn from Figures 14–17 affirm the aforementioned statement. Whether in terms of final precision or model convergence speed, none of the other models could parallel the remarkable performance of the complete FDFE module.

4.3. Loss Function Evaluation

The loss function is an essential component of a deep learning model because it tells the model how to measure the difference between the predicted and actual results and can reflect the convergence of the model during the network training process. In this study, the loss functions of five network models were compared. Figure 18 shows the loss curves for the training set, and Figure 19 shows the loss curves for the validation set.

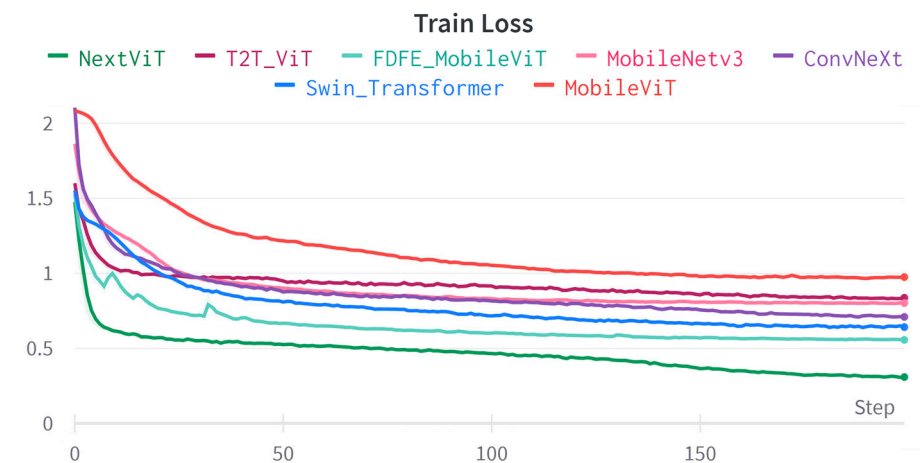


Figure 18. Training loss curves.

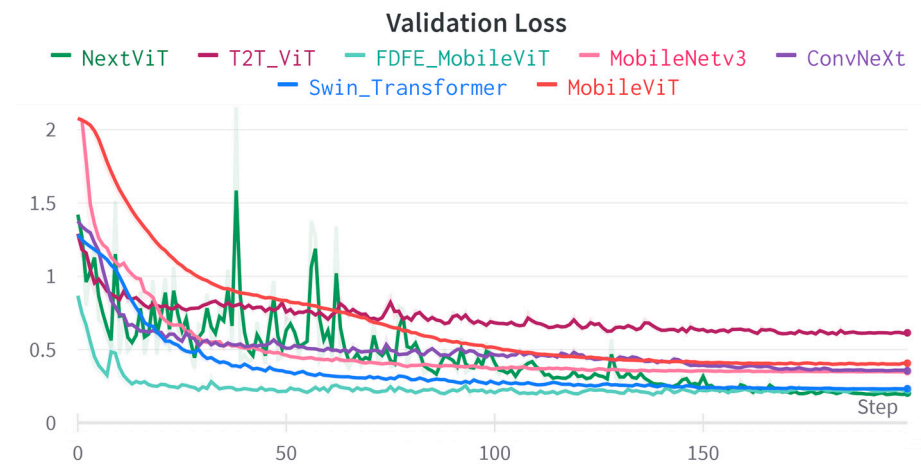


Figure 19. Validation loss curves.

Figure 18 shows that the convergence of FDFE-MobileViT was faster, and it converged near the 60th round of model iteration. Figure 19 shows that the model also converged around the 60th round with the validation set, and it converged faster than the other networks.

4.4. Confusion Matrix Comparison

Figure 20a–g show the confusion matrixes for FDFE-MobileViT, MobileViT, MobileNetv3, ConvNeXt, Swin-Transformer, NextViT, and T2T-ViT, where each row of the matrixes represents the actual category and each column represents the category predicted by the model. The diagonal elements of the matrixes represent the categories that the model correctly predicted. According to the confusion matrixes, the FDFE-MobileViT model could predict eight categories more accurately.

In the indoor static environment, FDFE-MobileViT had 32 indoor legal channels predicted as illegal channels and 13 illegal indoor channels predicted as legal channels. In comparison, the unimproved MobileViT had 104 illegal channels predicted as legal channels and 44 legal channels predicted as illegal channels in the indoor case. The comparison algorithm used in ConvNeXt was the best. In the dynamic indoor environment, the comparison algorithms used in ConvNeXt, Swin-Transformer, and MobileNetv3 were all more effective. Nevertheless, FDFE-MobileViT was in the lead, with fewer prediction errors than the above three algorithms.

By comparing the channel feature matrixes in the indoor environment with the generated feature maps, it was found that the difference between the channel feature matrixes for the legal and illegal channels in the indoor environment was manageable. Observing noticeable differences in the generated maps of the channel feature matrixes using the naked eye was challenging. The similarity of several classes was very high, which could have led to the model's inability to distinguish the difference between classes in the indoor case during the extraction of features. However, the total number of prediction errors remained low and acceptable.

In the static outdoor environment, FDFE-MobileViT did not predict illegal channels as legal channels or legal channels as illegal channels. In contrast, MobileViT predicted eight illegal channels as legal channels and seven illegal channels as legal channels. Similarly, we found slight differences between the outdoor legal and illegal channels by looking at the data collected in the outdoor environment. The generated feature maps were still indistinguishable with the naked eye, but the outdoor environment was more variable than the stable indoor environment, with vehicles and pedestrians passing by during the outdoor data collection process and signal blockage from plants also changing from time to time due to the outdoor environment. These effects were reflected in the collected data, so the data from the outdoor environment were more different than those from the indoor

environment. Overall, the models proposed in this study can perform excellently in the channel authentication task in indoor and outdoor environments.

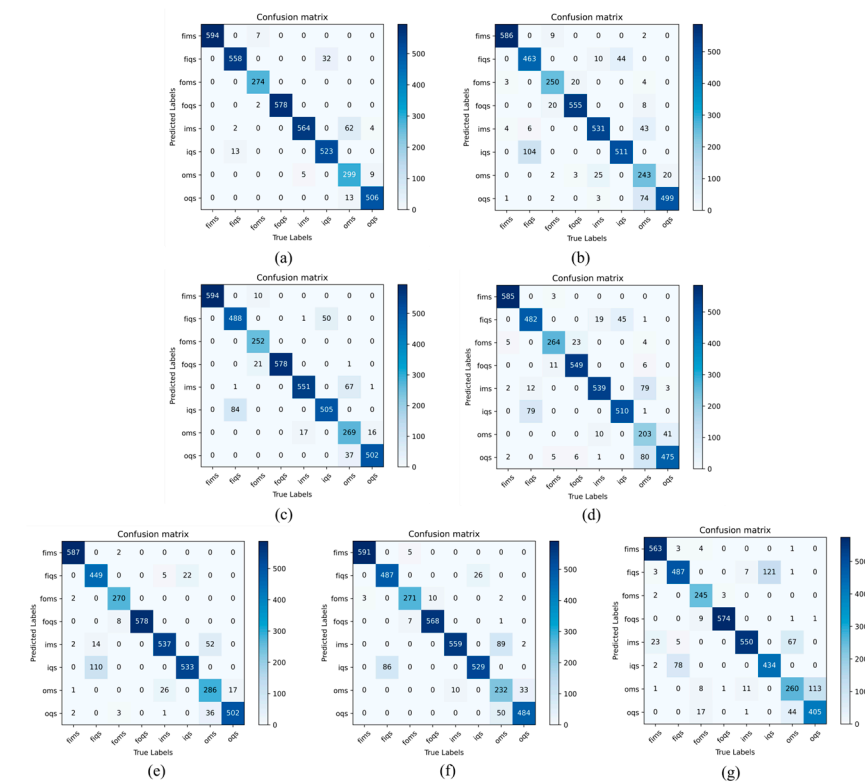


Figure 20. Confusion matrixes: (a) FDFE-MobileViT, (b) MobileViT, (c) MobileNetv3, (d) ConvNeXt, (e) Swin-Transformer, (f) NextViT, (g) T2T-ViT.

5. Conclusions

In the agricultural environment, cost is the primary consideration. Chen et al. [35] introduced a cluster-based physical-layer authentication algorithm that integrates clustering and lightweight symmetric cryptography with channel state information to enhance authentication accuracy. The research assumed that the terminal servers are high-performance, while devices in agricultural environments typically possess limited arithmetic power and may not meet the required computational demands. In contrast, the lightweight model proposed in this study necessitates fewer computational resources for the device and avoids the use of cryptographic assistance, thus making it more compatible with agricultural devices. Xiao et al. [36] proposed a scheme based on machine learning (ML) to achieve threshold-free physical-layer authentication across multiple landmarks. This approach can achieve higher authentication accuracy and reduced costs, but it requires the deployment of a significant number of peripheral devices, which is not cost-effective for agricultural environments. In contrast, our study’s model does not require additional devices, resulting in lower costs.

The training data also influence the contexts in which the model can be applied. Liao et al. [37] introduced a data augmentation-based multi-user PHY layer authentication scheme. Similarly to our study, their research utilized dynamic and static channel data from real-world factory environments. However, there are significant domain differences between industrial and agricultural data. Networks trained on industrial datasets may not be suitable for agricultural settings. Our model, developed using self-collected agricultural environment data, is better suited for such contexts. This study concatenated raw data to form multidimensional training and validation data. We processed the raw data into RGB images, avoiding redundancy while facilitating convenient input to various networks. While Liao et al. employed a more conventional DNN model, we utilized the lightweight

MobileViT model as a foundation and improved and optimized it for signal features, achieving superior authentication performance.

Most models seldom utilize the frequency-domain information for signal data compared to image data. However, the frequency-domain information for the signal contains the frequency component, frequency-domain amplitude, phase, and other signal information; this study proposed an FDFE module and used it to improve the MobileViT model. The FDFE module extracts the feature information in the signal's frequency domain to obtain more features, providing more information for the model to learn and remove in addition to the color. In this study, the authentication effect of FDFE-MobileViT was experimentally verified in natural smart agriculture indoor and outdoor environments. This study compared FDFE-MobileViT with Swin-Transformer, NextViT, T2T-ViT, Mobilenetv3, ConvNeXt, and MobileViT. A side-by-side comparison was conducted. FDFE-MobileViT achieved the best accuracy among the seven models at 96.6% with a small volume. The model proposed in this study may encounter the following challenges when deployed in large-scale agricultural settings. The diverse nature of large-scale agricultural environments, encompassing varying types of fields, crops, geographical conditions, and weather scenarios, contributes to data heterogeneity. Collecting extensive and diverse datasets at a large scale could present a challenge, particularly when acquiring data across different seasons and locations. In agricultural contexts, many applications necessitate real-time responsiveness and low latency, such as autonomous machinery and live monitoring. Addressing these requirements may involve optimizations in both model design and deployment strategies. In expansive agricultural landscapes, resource availability for hardware deployment could be restricted. Deep learning models typically demand substantial computational resources and storage capacity. Consequently, devising approaches to deploy models within resource-constrained environments is crucial, often entailing model compression, lightweight architectures, and acceleration techniques.

This study enhances the security of the agricultural IoT by combining MobileViT with the FDFE module. This amalgamation effectively strengthens the security of the agricultural Internet of Things, enabling more efficient protection of critical data and communication within the smart agriculture environment. Due to the lightweight nature of FDFE-MobileViT, it is more easily deployable with agricultural equipment, facilitating secure and dependable data transmission to elevate agricultural efficiency. Moreover, this research holds significant managerial implications. It provides agricultural managers with a foundational framework for making informed decisions and implementing security measures, enabling them to better utilize limited resources and enhance management efficiency.

This study centers on channel authentication within smart agricultural environments. During the data collection phase, we focused solely on a single area of crop farmland. As a result, other types of farmland, such as tea fields and fruit orchards, were not included in our dataset. In our future research endeavors, we will expand our data collection efforts to encompass a broader range of agricultural environments. This expansion will allow us to extend our model's applicability to more intricate settings. Additionally, we are considering the incorporation of more advanced backbone networks to enhance the overall performance of our model. Furthermore, we are exploring the possibility of adapting these advanced backbone networks to further optimize our model's performance. Looking ahead, we intend to extend the methodology employed in this study to various other environments, including factory and office settings. This approach aligns with our commitment to maintaining consistency in our research focus and broadening the scope of our methodologies.

Author Contributions: Conceptualization, F.P. and B.Z.; Data organization, B.Z., X.Z., L.S. and P.C.; Formal analysis, X.D. and F.P.; Funding acquisition, X.D.; Investigation, B.Z. and X.Z.; Methodology, F.P., B.Z. and X.Z.; Project management, F.P., X.D. and B.Z.; Resourcing, F.P. and B.Z.; Visualization, B.Z.; Writing—original draft, F.P. and B.Z.; Resources, F.P. and B.Z.; Validation, B.Z. and X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data available on request due to restrictions; e.g., privacy or ethical. The data presented in this study are available on request from the corresponding author.

Acknowledgments: Thanks to Xiaonuo Gou, Xueqin Jiang, and Gang Liu for their advice throughout the research process. Thanks to all the partners of AI Studio for their support. Thanks to the support of Ya'an Digital Agriculture Engineering and Technology Research Center and Sichuan Agricultural University's "Double Support Program".

Conflicts of Interest: The authors declare that they have no known conflicting financial interest or personal relationship that could have appeared to influence the work reported in this paper.

Abbreviations

| | |
|----------|--|
| Abridge | Interpretations |
| FDFE | Frequency-domain feature extraction |
| OFDM | Orthogonal frequency division multiplexing |
| MIMO | Multiple in multiple out |
| USRP | Universal software radio peripheral |
| Tx | Transport |
| Rx | Receive |
| ViT | Vision Transformer |
| CPU | Central processing unit |
| GPU | Graphics processing unit |
| FFT | Fast Fourier transform |
| DSC Conv | Depthwise separable convolution |
| Conv | Convolution |
| MV2 | Inverted residual block structure in Mobilenetv2 |
| SGD | Stochastic gradient descent |
| ICBAW | Illegal channel before adding windows |
| ICAAW | Illegal channel after adding windows |
| LCBAW | Legal channel before adding windows |
| LCAAW | Legal channel after adding windows |

Notation

| | |
|-------------|--|
| Y | Received data |
| X | Sent data |
| H | Channel state information |
| b | Noise data |
| X_T | Input tensor of the FDFE module |
| e | Euler number |
| j | Imaginary unit |
| X_F | Frequency-domain tensor corresponding to X_T |
| W_n | Set of Bartlett windows |
| X_w | Result for X_F after the windowing operation |
| K_c | Convolution kernel |
| X_c | Intermediate variable |
| P_c | Convolution kernel |
| Y_F | Output tensor of DSC Conv |
| Y_A | The output result of the FDFE module |
| y_i | i th sample |
| \hat{y}_i | Prediction probability of the i th sample |
| lr | Current learning rate |
| lrf | Final decay rate of the learning rate |
| lri | Initial learning rate |
| T_{cur} | Current number of training rounds |
| T_{max} | Total number of training rounds |

References

1. Shen, Y. Construction of a Wireless Sensing Network System for Leisure Agriculture for Cloud-Based Agricultural Internet of Things. *J. Sens.* **2021**, *2021*, 3021771. [[CrossRef](#)]
2. Li, G.; Li, D.; Chen, W.; Zhang, Y.; Xu, S. Design and Application of Special Sensors and Internet of Things (IoT)-Based Wireless System for Agricultural Information Monitor. *J. Phys. Conf. Ser.* **2020**, *1646*, 012130. [[CrossRef](#)]
3. Kim, W.-S.; Lee, W.-S.; Kim, Y.-J. A Review of the Applications of the Internet of Things (IoT) for Agricultural Automation. *J. Biosyst. Eng.* **2020**, *45*, 385–400. [[CrossRef](#)]
4. He, M. Research on Rural Water-Saving Intelligent Irrigation System Based on Internet of Things. In Proceedings of the 2022 International Conference on Automation, Robotics and Computer Engineering (ICARCE), Virtual, 16 December 2022; pp. 1–3. [[CrossRef](#)]
5. Tsang, Y.P.; Wu, C.H.; Ip, W.H.; Shiau, W.L. Exploring the Intellectual Cores of the Blockchain–Internet of Things (BIoT). *J. Enterp. Inf. Manag.* **2021**, *34*, 1287–1317. [[CrossRef](#)]
6. Raya, M.; Hubaux, J.-P. Securing vehicular ad hoc networks. *J. Comput. Secur.* **2007**, *15*, 39–68. [[CrossRef](#)]
7. Biswas, S.; Mišić, J. A Cross-Layer Approach to Privacy-Preserving Authentication in WAVE-Enabled VANETs. *IEEE Trans. Veh. Technol.* **2013**, *62*, 2182–2192. [[CrossRef](#)]
8. Asim, M.; Ignatenko, T.; Petkovic, M. Hierarchical Attribute-Based Encryption and Decryption. U.S. Patent 10,211,984, 19 February 2019.
9. Tsai, K.-L.; Huang, Y.-L.; Leu, F.-Y.; You, I.; Huang, Y.-L.; Tsai, C.-H. AES-128 Based Secure Low Power Communication for LoRaWAN IoT Environments. *IEEE Access* **2018**, *6*, 45325–45334. [[CrossRef](#)]
10. Jakes, W.C.; Cox, D.C. *Microwave Mobile Communications*; Wiley-IEEE Press: Hoboken, NJ, USA, 1994; ISBN 0-7803-1069-1.
11. Shafiee, S.; Ulukus, S. Achievable Rates in Gaussian MISO Channels with Secrecy Constraints. In Proceedings of the 2007 IEEE International Symposium on Information Theory, Nice, France, 24–29 June 2007; pp. 2466–2470.
12. Marabissi, D.; Mucchi, L.; Stomaci, A. IoT Nodes Authentication and ID Spoofing Detection Based on Joint Use of Physical Layer Security and Machine Learning. *Future Internet* **2022**, *14*, 61. [[CrossRef](#)]
13. Zhang, H.; Ma, C.; Pazzi, V.; Li, T.; Casagli, N. Deep Convolutional Neural Network for Microseismic Signal Detection and Classification. *Pure Appl. Geophys.* **2020**, *177*, 5781–5797. [[CrossRef](#)]
14. Blanc, F.; Syed, A.; Esfahani, A.M.; Venna, S.R.; Ajila, S.A. A Deep Learning Sequential-Based Model for Predicting Victories in Video Games. In Proceedings of the 2022 IEEE 23rd International Conference on Information Reuse and Integration for Data Science (IRI), San Diego, CA, USA, 9–11 August 2022; pp. 212–213.
15. Ito, T.; Maeno, T.; Tsuchikame, H.; Shishido, M.; Nishi, K.; Kojima, S.; Hayashi, T.; Suzuki, K. Adapting a Low-Count Acquisition of the Bone Scintigraphy Using Deep Denoising Super-Resolution Convolutional Neural Network. *Phys. Med. Eur. J. Med. Phys.* **2022**, *100*, 18–25. [[CrossRef](#)]
16. Hu, S.; Shuai, L.; Yang, Q.; Chen, H. Study on Wireless Signal Propagation in Residential Outdoor Activity Area Based on Deep Learning. In Proceedings of the 2021 International Conference on Computer, Control and Robotics (ICCCR), Shanghai, China, 8–10 January 2021; pp. 225–230.
17. Yadav, K.; Kariri, E.; Alotaibi, S.D.; Viriyasitavat, W.; Dhiman, G.; Kaur, A. Privacy Protection against Attack Scenario of Federated Learning Using Internet of Things. *Enterp. Inf. Syst.* **2023**, *17*, 2101025. [[CrossRef](#)]
18. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In *Proceedings of the Advances in Neural Information Processing Systems*; Curran Associates, Inc.: New York, NY, USA, 2017; Volume 30.
19. Mehta, S.; Rastegari, M. MobileViT: Light-Weight, General-Purpose, and Mobile-Friendly Vision Transformer. *arXiv* **2021**, arXiv:2110.02178.
20. Massey, J.W.; Starr, J.; Lee, S.; Lee, D.; Gerstlauer, A.; Heath, R.W. Implementation of a Real-Time Wireless Interference Alignment Network. In Proceedings of the 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, CA, USA, 4–7 November 2012; pp. 104–108.
21. Proposed Model for Sustainable and Scalable Vertical Farm. *Nile J. Commun. Comput. Sci.* **2022**, *2*, 17–28. [[CrossRef](#)]
22. Fu, J.; Juyal, P.; Zajić, A. 300 GHz Channel Characterization of Chip-to-Chip Communication in Metal Enclosure. In Proceedings of the 2019 13th European Conference on Antennas and Propagation (EuCAP), Krakow, Poland, 31 March–5 April 2019; pp. 1–5.
23. Zhao, S.J. Research on the shielding of wireless signals by trees. *Sci. Technol. Inf.* **2007**, *2*, 38–39. [[CrossRef](#)]
24. Giacomini, J.C.; Vasconcelos, F.H.; da Silva, E.J. Estimating Vegetation Water Content with Wireless Sensor Network Communication Signals. In Proceedings of the 2007 IEEE Instrumentation & Measurement Technology Conference IMTC 2007, Warsaw, Poland, 1–3 May 2007; pp. 1–5.
25. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
26. Hou, Q.; Zhou, D.; Feng, J. Coordinate Attention for Efficient Mobile Network Design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13713–13722.
27. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer: Cham, Switzerland, 2018; pp. 3–19.
28. Liu, Y.; Shao, Z.; Hoffmann, N. Global Attention Mechanism: Retain Information to Enhance Channel-Spatial Interactions. *arXiv* **2021**, arXiv:2112.05561.

29. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv* **2017**, arXiv:1704.04861.
30. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 10012–10022.
31. Li, J.; Xia, X.; Li, W.; Li, H.; Wang, X.; Xiao, X.; Wang, R.; Zheng, M.; Pan, X. Next-ViT: Next Generation Vision Transformer for Efficient Deployment in Realistic Industrial Scenarios. *arXiv* **2022**, arXiv:2207.05501.
32. Yuan, L.; Chen, Y.; Wang, T.; Yu, W.; Shi, Y.; Jiang, Z.-H.; Tay, F.E.H.; Feng, J.; Yan, S. Tokens-to-Token ViT: Training Vision Transformers From Scratch on ImageNet. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 558–567.
33. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for MobileNetV3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
34. Liu, Z.; Mao, H.; Wu, C.-Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A ConvNet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 11976–11986.
35. Chen, Y.; Wen, H.; Wu, J.; Song, H.; Xu, A.; Jiang, Y.; Zhang, T.; Wang, Z. Clustering Based Physical-Layer Authentication in Edge Computing Systems with Asymmetric Resources. *Sensors* **2019**, *19*, 1926. [[CrossRef](#)] [[PubMed](#)]
36. Xiao, L.; Wan, X.; Han, Z. PHY-Layer Authentication with Multiple Landmarks With Reduced Overhead. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 1676–1687. [[CrossRef](#)]
37. Liao, R.-F.; Wen, H.; Chen, S.; Xie, F.; Pan, F.; Tang, J.; Song, H. Multiuser Physical Layer Authentication in Internet of Things With Data Augmentation. *IEEE Internet Things J.* **2020**, *7*, 2077–2088. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.