*electronics*

*Article*

# Hierarchical Security Authentication with Attention-Enhanced Convolutional Network for Internet of Things

**Xiaoying Qiu [1,*], Guangxu Zhao [1], Jinwei Yu [2], Wenbao Jiang [1], Zhaozhong Guo [1] and Maozhi Xu [3]**

[1] College of Computer Science, Beijing Information Science and Technology University, Beijing 100192, China; 2023020981@bistu.edu.cn (G.Z.); 20051452@bistu.edu.cn (W.J.); guozhaozhong@bistu.edu.cn (Z.G.)
[2] School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; hahayuswag@bupt.cn
[3] School of Mathematical Sciences, Peking University, Beijing 100871, China; mzxu@math.pku.edu.cn
* Correspondence: 20192329@bistu.edu.cn

**Abstract:** As security authentication issues continue to arise in future wireless communication networks, researchers are working hard to further improve authentication techniques. Recently, physical layer authentication (PLA) has received widespread attention for its lightweight nature compared to traditional encryption methods based on keys and blockchain. However, the existing PLA mechanisms based on a fixed decision threshold have low reliability in dynamic environments. Moreover, PLA solutions are typically based on binary authentication, and these binary-type schemes cannot provide different levels of access control. To address these challenges, this article introduces the concept of hierarchical security authentication, aiming to achieve multi-level secure authorization access. In order to further improve the accuracy of identity verification, we design an Attention-Enhanced Convolutional Network (AECN) model that integrates the attention mechanism. Specifically, by introducing a confidence score branch, the proposed AECN-based PLA scheme completes authentication without a threshold, thus avoiding the issues stemming from inappropriate threshold settings in conventional PLA schemes. The simulation results show that our proposed AECN framework outperforms existing algorithms at different levels of security authentication.

**Keywords:** artificial intelligence; blockchain security; physical layer security; privacy protection; machine learning

check for updates

**Citation:** Qiu, X.; Zhao, G.; Yu, J.; Jiang, W.; Guo, Z.; Xu, M. Hierarchical Security Authentication with Attention-Enhanced Convolutional Network for Internet of Things. *Electronics* **2024**, *13*, 4699. https://doi.org/10.3390/electronics13234699

Academic Editor: Stefano Scanzio

Received: 31 October 2024
Revised: 22 November 2024
Accepted: 26 November 2024
Published: 28 November 2024

**Copyright:** © 2024 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

The Internet of Things (IoT), as the core technology of the fourth industrial revolution [1], has become a significant driver of global scientific and technological advancement, as well as industrial modernization. It promotes technological advancements in fields such as smart healthcare [2], smart home [3] and wearable devices [4], resulting in significant improvements in work and lifestyle [5]. However, the large number of connected smart devices in IoT networks make them vulnerable to spoofing attacks, making it particularly important to prevent unauthorized access to resources [6]. The conventional authentication procedures predominantly rely on upper-layer protocols [7] and cryptographic algorithms [8]. Blockchain, a distributed ledger technology, is used in cryptographic algorithms for its secure, trustworthy and tamper-proof feature [9]. For example, in [10], blockchain is used to store key ciphertext. The deployment of key-based encryption technology for repeated authentication in latency-sensitive IoT environments may result in latency and computational overhead, posing a challenge to the processing capabilities of these smart IoT devices [11]. In addition, the rapid development of quantum computing has also posed challenges to system security, especially in brute force password cracking scenarios. In view of this, more efforts are needed to enhance the security authentication technology of IoT.

In recent years, physical layer authentication (PLA) [12] has become a novel security authentication scheme that utilizes the spatiotemporal uncorrelated characteristics of different channels to achieve identity recognition. As is supported by the well-known Jakes model, the received signals decorrelate over a distance of half a wavelength. It is difficult for an attacker to forge the channel between legitimate transceivers [13]. Upon receiving the signal, the receiver verifies its legitimacy by confirming whether it originates from a permitted or unauthorized device. This process not only ensures the validity of the signal but also traces its source, effectively identifying unauthorized users. Common physical layer attributes include radio frequency (RF) fingerprinting [14], channel information [15], and carrier frequency offset (CFO) [16], which are difficult for attackers to imitate. However, RF fingerprinting techniques face two different forms of replay attacks: signal replay and feature replay. Furthermore, RF features are susceptible to fluctuations caused by changes in the channel environment [17]. Therefore, channel state information (CSI) is employed for the secure authentication of devices accessing the IoT, and the reasons are outlined as follows: In wireless communication, CSI delineates the characteristics of signal propagation channels, encompassing signal attenuation, phase offset, multipath effect, and more. Distinct physical conditions between transmitters and receivers give rise to variances in CSI, endowing it with notable variability and distinctiveness, which renders CSI a potent identifier for authentication purposes. The resilience of CSI to alteration is a consequence of intrinsic dependence on physical environments. Consequently, malevolent entities encounter considerable challenges when attempting to falsify or imitate authentic CSI with the intention of subverting authentication systems.

Research in PLA has concentrated on two principal areas: firstly, statistical model-based PLA, which employs parameters or thresholds to distinguish between different users; and secondly, PLA using machine learning (ML) and deep learning (DL), whereby network models are trained to make autonomous decisions.

**Statistical model-based PLA:** Conventional PLA relies on hypothesis testing, which typically involves setting a threshold to complete the authentication process [12]. Xiao et al. [18] proposed a statistical distribution model that incorporates established or estimated parameters to characterize channel attribute. Subsequently, hypothesis testing was employed to define a precise threshold. The proposed approach leverages the inherent similarities and differences in channel characteristics as a means of discerning legitimate users from potential intruders. In [19], Lu considered the influence of general channel conditions in the system model, proposed a scheme to improve authentication performance by exploiting channel correlation coefficient information, and provided closed-form expressions for decision thresholds under binary hypothesis testing. Xie et al. [20] presented two PLA schemes and performed a theoretical analysis to derive the closure expressions. The results of the subsequent experimental investigation demonstrate an improvement in authentication performance. Due to the dynamic changes in channel conditions, traditional hypothesis testing methods that rely on thresholds are no longer applicable. Xiao et al. showed that, through reinforcement learning technology, even when the system parameters are unknown, the threshold can be dynamically adjusted to optimize the detection performance. However, despite the optimal adjustment of the threshold, limitations persist in the detection performance [21,22]. Blockchain is used to assist security authentication due to its decentralization and tamper-proof quality [23]. Fang et al. [24] proposed a blockchain-assisted handover authentication scheme for the Internet of Vehicles (IoV), utilizing the threshold for authentication and storing information on the blockchain.

**ML and DL-based PLA:** Classical classification algorithms, in particular support vector machine (SVM), are used in PLA. Pei et al. [25] employed SVM to construct classifiers that could effectively differentiate between channel feature vectors with high and low similarity. SVM technology is implemented for user authentication at the packet level, ensuring the ability to distinguish different users even in cases where their signal fingerprints exhibit similarity [26]. In [27], a multi-dimensional adaptive PLA scheme is presented, which employs fuzzy theory to model multiple physical layer attributes that are inherently

imperfect and uncertain, thereby enabling the design of a compact authentication model. In [28], a collaborative PLA scheme based on horizontal federated learning is proposed, which distributes computational tasks among multiple edge nodes or cloud servers. Nevertheless, the preliminary authentication process does not occur at the physical layer. In order to address the challenge of determining testing thresholds in dynamic network environments, Wang et al. [29] used the multi-dimensional characteristics of wireless channels and training data generated from pseudo-spoofing models. In [30], a threshold-free PLA approach is proposed as an alternative to traditional threshold-based decision-making. In contrast, the channel matrix estimated from wireless nodes is employed as the authentication input, thereby facilitating more adaptive classification through the application of ML algorithms. Qiu et al. [31] proposed a PLA enhancement scheme based on a Gaussian mixture model (GMM), which conducted clustering analysis on channel information through a two-dimensional feature space to detect and distinguish legitimate users from potential spoofing attackers. A PLA scheme based on logistic regression is proposed [32]. The training of the model is expedited by reducing the training dataset [33] and optimizing the training process [34], while ensuring that the authentication performance remains uncompromised. In [33], Meng demonstrated that authentication performance can be optimized even with a limited training dataset by introducing Gaussian noise into the smooth potential space. By integrating a deep neural network (DNN), a multi-user PLA scheme with data augmentation is proposed, and three gradient descent methods are proposed as a means of accelerating the training of the network [34]. Nevertheless, the above scheme adopts binary authentication, which allows the receiver to either accept or reject. Consequently, a binary authentication scheme results in misdetections, leading to a waste of resources.

In short, it is difficult for the PLA to set appropriate thresholds when channel conditions dynamically change, as fixed thresholds not adapt to rapid changes, leading to misjudgments or missed detections. Moreover, existing binary authentication schemes are unable to provide multi-level access control for the system. Therefore, we propose a hierarchical security authenticator for PLA, which includes a confidence branch and an attention-enhanced convolutional network (AECN). The proposed AECN model, by introducing hierarchical security authenticator, enables multi-level secure authorized access, thereby better accommodating the needs of different devices and users in IoT environments. The specific contributions of this work are summarized below:

- To provide multi-levels access control, a hierarchical security authenticator is proposed. The receiver evaluates the trustworthiness of the smart devices based on the proposed PLA solution. In our proposed neural networks, we transform the hierarchical security issue into a multi-classification problem for security authentication.
- A confidence score branch is introduced, which is designed to evaluate the model's prediction probability and assign a corresponding confidence score. Furthermore, we propose an attention module that incorporates a bottleneck structure and depth-wise separable convolution. The proposed AECN model integrates an attention module that processes the input feature map, generates attention weights, and allows the model to focus on key information, thereby enhancing the accuracy of identity verification.
- We conducted comparative experiments on our proposed AECN scheme and CNN-based PLA scheme. The simulation results show that our solution provides more reliable protection for IoT devices at different levels of security authentication.

The subsequent sections of this article are structured as follows: Section 2 offers a comprehensive overview of the system model. Section 3 introduces the proposed hierarchical security authentication scheme, hierarchical security authenticator, and attention module. Section 4 is the presentation of evaluation indicators and performance analysis. Section 5 is the discussion and future work. Section 6 is the conclusions. Abbreviations lists the abbreviations that appear in this paper and their spelled-out terms.

## 2. System Model

As shown in Figure 1, the classic Alice–Bob–Eve communication model is considered., in which IoT nodes ($\text{Alice}_1$, $\text{Alice}_2$, …, $\text{Alice}_N$) carry out secure communication with the base station (BS) Bob. The communication between $\text{Alice}_i$ and Bob requires secure and continuous signal transmission. Bob needs to verify the received message to ascertain whether it was sent by $\text{Alice}_i$. Eve is an illegitimate sending node, also known as a spoofer. Eve attempts to mimic the identity of a legitimate device $\text{Alice}_i$, conducting deception through an independent spoofing channel to gain unauthorized access to the system or introduce false data. Spoofing attacks lead to unauthorized access, privacy breaches, and erroneous decisions in IoT scenarios.

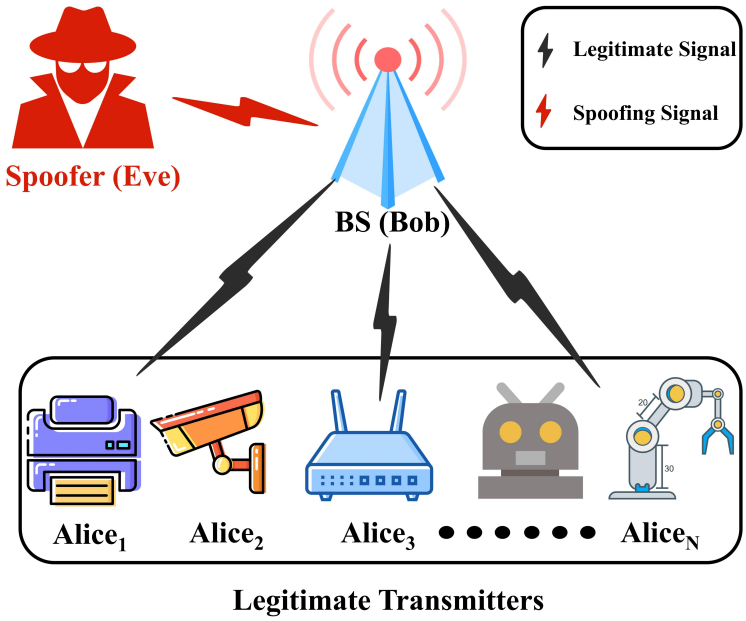


**Figure 1.** System model.

Assuming Bob is furnished with $N_\text{r}$ antennas and $\text{Alice}_i$ has $N_\text{t}$ antennas, it is assumed that Orthogonal Frequency Division Multiplexing (OFDM) technology is used in the communication process, and $N_\text{s}$ subcarriers are allocated for data transmission. Then, we have

$$\overrightarrow{P} = \mathbf{CSI} \cdot \overrightarrow{Q} + \overrightarrow{N} \tag{1}$$

where $\overrightarrow{P}$ and $\overrightarrow{Q}$ represent the received signal vectors and the transmitted signal vectors, $\overrightarrow{N}$ denotes the Gaussian noise, and **CSI** represents the channel state information matrix. The matrix $\mathbf{CSI}_\text{mn}$ is a complex value. Defined as

$$\mathbf{CSI}_\text{mn} = \text{real}(\mathbf{CSI}_\text{mn}) + \text{img}(\mathbf{CSI}_\text{mn}) \tag{2}$$

where the real value represents the $m$-th antenna and the imaginary value represents the $n$-th subcarrier.

$\text{Alice}_i$ and Bob utilize the unique physical layer attribute as the basis for authentication and have established an initial authentication mechanism before the communication begins, which serves as a prerequisite for PLA. At initial authentication stage, the estimation value of $\text{Alice}_i$'s CSI can be obtained, which is represented as

$$C_{A1}, C_{A2}, \ldots, C_{At} \tag{3}$$

where $C_{At}$, with $t \in \{1, 2, \ldots, T\}$, represents the information received from a legitimate device. The new CSI estimation $C_{Ot}$ $(t = 1, 2, \ldots)$ clearly shows the examples of physical

layer estimation. The new CSI estimation $C_{Ot}$ may come from Alice$_i$ or Eve. The CSI at $t + 1$ is represented as

$$\Delta C_{O1} = f(C_{A1}, C_{A2}, \ldots, C_{AT}, C_{O1}) \quad (4)$$

where $f$ represents the discrepancy evaluation function between $C_{A1}$, $C_{A2}$, …, $C_{AI}$ and $C_{O1}$. Furthermore, a hierarchical security authentication model is introduced, comprising $n$ categories, with each category assigned a corresponding authorization level based on the different resources.

## 3. Hierarchical Security Authentication Scheme and Attention Module

We propose a hierarchical security authenticator for PLA that includes a confidence score branch and an attention module. The confidence score branch is responsible for authenticating the receiver by assigning confidence score to the predicted probabilities, and the attention module generates attention weight.

### *3.1. Overall Framework of Hierarchical Security Authentication Scheme*

Figure 2 shows the overall framework of Hierarchical Security Authentication scheme for multi-level access control of IoT devices. During the data collection phase, BS obtains the original CSI values from the IoT devices. In the stage of converting CSI data to images, the obtained CSI data are constructed as images as input to the authenticator. The CSI matrix is considered in the form of an image, where the first channel is a real value and the second channel is a imaginary value. Considering that there are 64 antennas and 52 subcarriers in the IoT scene, the CSI matrix can be viewed as a three-dimensional image with dimensions of 64 × 52 × 2. Afterwards, the CSI image samples are input into the authenticator for training to obtain training weight. In the course of the online authentication phase, BS verifies the received data packets by making use of the trained authenticator and the recently obtained CSI image. The authenticator furnishes a prediction probability and security score, and the purpose is to ascertain the legitimacy of the data packet's source device.

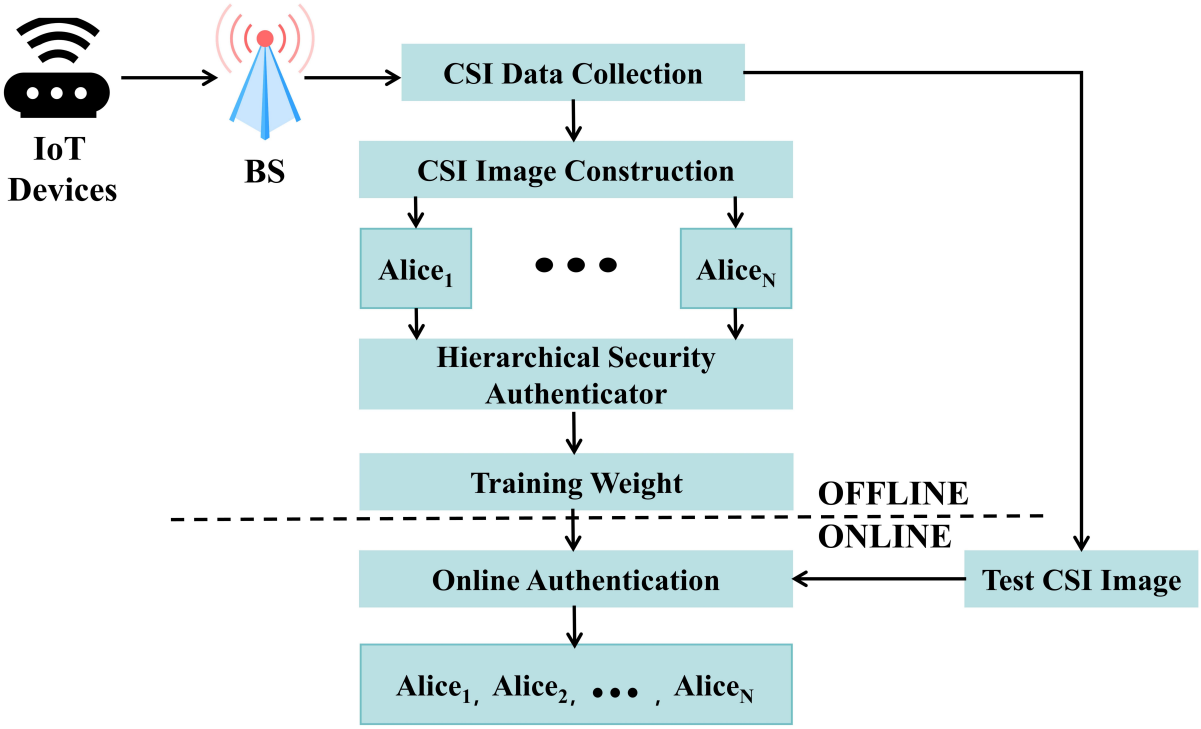


**Figure 2.** Overall framework of hierarchical security authentication scheme.

### 3.2. Hierarchical Security Authenticator

Figure 3 illustrates the comprehensive architectural framework of our proposed Hierarchical Security Authenticator, which integrates our proposed AECN and confidence score branch. During the data processing phase, the received signals are used to extract CSI attributes, which are then processed as a training dataset. Subsequently, the data dimensions are adjusted to a four-dimensional matrix of $64 \times 52 \times 2 \times 40{,}000$, where 40,000 is the number of CSI images, before being input to the AECN.

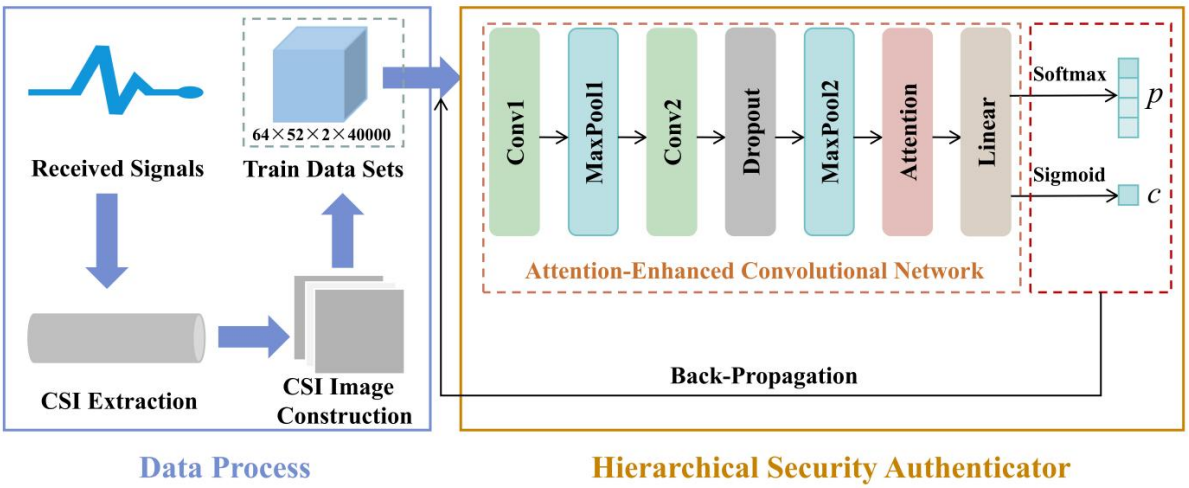


**Figure 3.** Architecture of Hierarchical Security Authenticator.

The AECN model comprises a series of layers that are designed to facilitate the extraction and classification of features. The network initially extracts preliminary features of the input CSI image through convolutional layer 1 (Conv1). Next, MaxPool1 is used to reduce the spatial dimension of the feature map and focus on preserving the most important features. Subsequently, the second convolutional layer (Conv2) serves to further refine the process of feature extraction. To prevent overfitting, a dropout layer is incorporated into the network. This is achieved by randomly discarding a portion of the activation values of neurons. MaxPool2 continues the process of dimensionality reduction, further abstracting the feature representations. The core of the network is the attention layer, which emphasizes the features in the input feature map through attention mechanisms, enhancing the recognition ability of the model. For a detailed description of this attention layer, please refer to Section 3.3. Subsequently, the linear layer transforms these features into a prediction space suitable for classification task.

We use AECN to automatically learn a representation of CSI feature maps. CSI images provide a structured form of data for the AECN model, facilitating the model to automatically learn complex spatial features through convolutional neural networks. This approach captures subtle changes in the CSI data and enhances the recognition ability of the model.

Both the branch tasked with calculating the confidence score and the one responsible for generating the Softmax predictions are linked to a same fully connected layer. The Softmax function is applied in order to obtain prediction probabilities $p$, while the sigmoid function is applied in order to obtain the confidence score $c$. The authenticator should output $c$ close to 1 if it is confident that it can produce a correct prediction for the given input. Conversely, if the authenticator is not confident that it can produce the correct prediction, it should output $c$ close to 0.

$$p, c = F(img, \theta), p_i, c \in [0, 1], \sum_{i=1}^{N} p_i = 1 \tag{5}$$

where *img* is the training data of the input convolutional network and $\theta$ denotes parameters of the AECN.

The softmax prediction probabilities are adjusted during training to provide authenticator with the hints they need. This is done by interpolating between the original predictions and the target probability distribution $y$, with the degree of interpolation indicated by the confidence.

$$p' = c \cdot p + (1 - c)y \tag{6}$$

A negative log-likelihood function is employed to calculate the task loss. However, to prevent the network from consistently selecting $c = 0$ to minimize the loss function, a logarithmic penalty mechanism is incorporated into the confidence loss. The revised loss function $L$ combines the task loss $L_{AECN}$ with the confidence loss $L_c$, weighted by the hyperparameter $\lambda$, in order to achieve a balance between the task loss and the confidence loss.

$$L = L_{AECN} + \lambda L_c = -\sum_{i=1}^{N} \log(p_i')y_i + -\lambda \log(c) \tag{7}$$

### *3.3. Attention Module*

The attention module is shown in Figure 4, which uses a bottleneck structure and depth-wise separable convolution to generate attention weights. The following presents the specific details:

The feature map is fed into a bottleneck structure, which aims to reduce the input dimensionality of the feature while preserving crucial information. The first component of the bottleneck structure is a convolutional layer with a kernel_size of 1, which independently processes the neighborhood of each pixel, thereby reducing the number of channels in the feature map. This is followed by a BatchNorm2d layer that accelerates the training process, and its output undergoes non-linear processing through a ReLU activation function. The feature map is processed by a depth-wise separable convolutional layer, which consists of a depth-wise convolutional layer and a pointwise convolutional layer. The depth-wise convolutional layer performs spatial convolution independently for each input channel, using a kernel_size of 3. Following this layer, there is another BatchNorm2d layer and a ReLU activation function. Then, the pointwise convolutional layer transforms the output of the depth-wise convolutional layer across channels. After the bottleneck structure, a second $1 \times 1$ convolutional layer is utilized to adjust the number of channels. Subsequently, the features undergo batch normalization for standardization and are activated by the ReLU function to introduce non-linearity. These processed features then pass through a Sigmoid activation function, which scales their values to the range of [0, 1], representing the attn_weights for each feature location. The computed attention weights attn_weights are utilized to enhance the feature representation of the input feature maps $x$. This is achieved by performing an element-wise multiplication between the attention weights and the original input feature maps:

$$\text{attended_features} = x \odot \text{attn_weights} \tag{8}$$

where $\odot$ represents the element-wise multiplication operation. By multiplying the input features with their corresponding attention weights, the model is able to focus more on important features and suppress less significant information. This process effectively enhances the discriminative power of the feature representation.

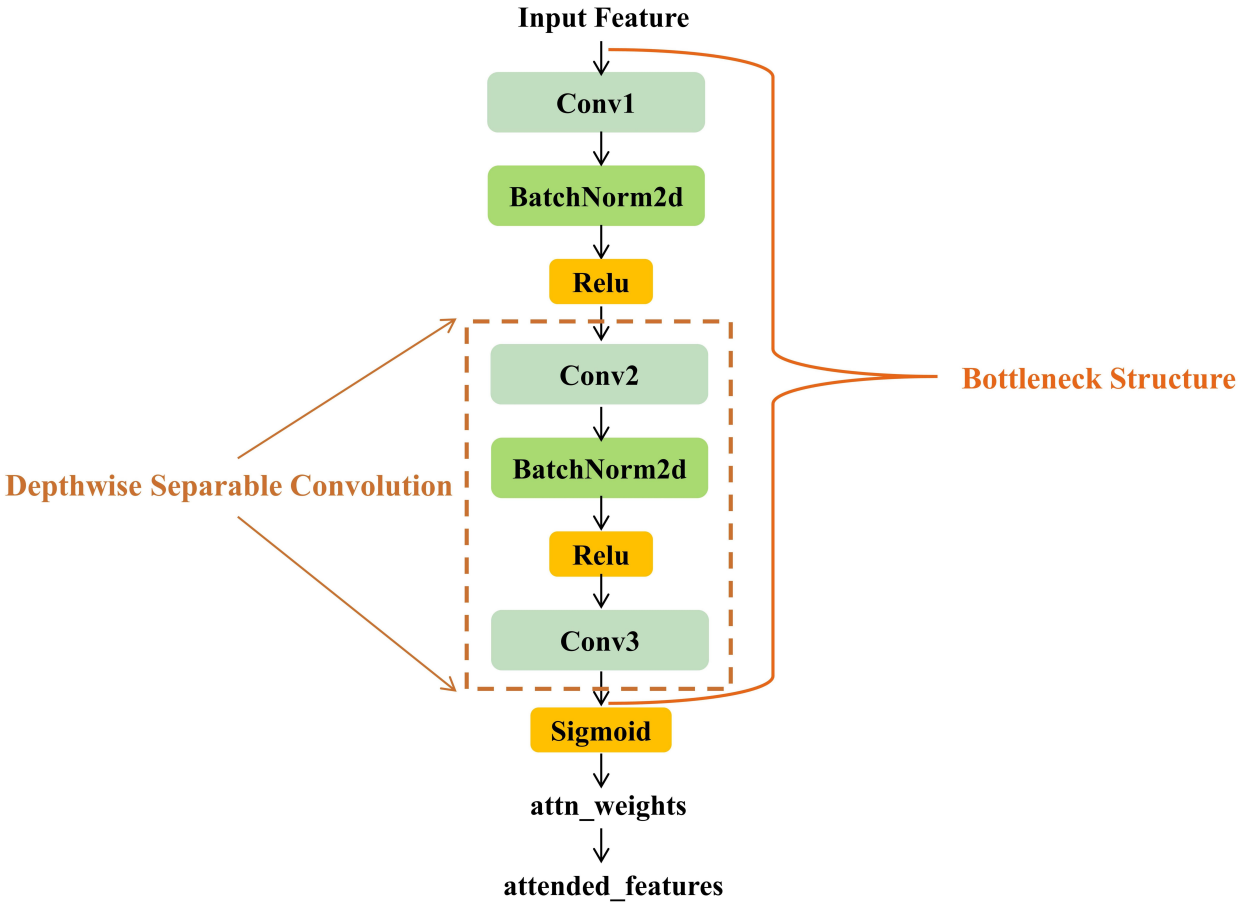


**Figure 4.** The proposed attention module.

## 4. Performance Evaluation

In this section, we compare the performance of the CNN with the AECN model, which integrates an attention mechanism. In this paper, we define two different security authentication levels, N = 5 and N = 10, to denote different levels of stringency in the authentication process to control the rights of a device to perform resource access (there are incremental security requirements between each level). The experimental environment is built on the PyTorch 2.1.0 framework, executed within a Python 3.10 environment, and hosted on Ubuntu 22.04 to ensure compatibility and stability. Furthermore, by leveraging CUDA 12.1 acceleration technology, we utilize a high-performance RTX 4090 GPU with 24 GB of memory to maximize training efficiency. The principal parameter values are presented in Table 1.

**Table 1.** Parameter values.

| Parameter | Value |
|---|---|
| Batch size | 180 |
| Optimizer | Stochastic Gradient Descent (SGD) |
| Inital learning rate | $1 \times 10^{-2}$ |
| Weight decay | $1 \times 10^{-3}$ |
| Momentum | $9 \times 10^{-1}$ |
| Nesterov | True |

### *4.1. Dataset and Implementation Trick*

We utilized MATLAB R2020a simulation to compute the signals from both the direct and reflection paths at various locations. The signal at each position is obtained by subtracting the direct and reflection paths to simulate multipath effects in actual signal propagation. In the IoT scenario, consider a BS with 64 antennas, each collecting CSI data from 52 various subcarriers. The CSI matrix is considered in the form of an image, where the first channel is a real value and the second channel is a imaginary value. Considering that there are 64 antennas and 52 subcarriers in the IoT scene, the CSI matrix can be viewed as a three-dimensional image with dimensions of $64 \times 52 \times 2$. The training data consists of 40,000 CSI images, and the testing data consist of 10,000.

In the course of our experiments, learning rate is one of the most important hyperparameters. In order to optimize the model performance and avoid overfitting, a learning rate strategy with decreasing segments is employed. In particular, the learning rate is decreased at the 6th, 12th, 18th, 24th, and 30th epochs, reducing to 20% of the original value at each stage. A range of learning rates was explored, commencing at 0.01 and decreasing sequentially to 0.002, 0.0004, 0.00008, 0.000016, and 0.0000032. This exponentially decaying learning rate tuning strategy has been demonstrated to be effective in deep learning training.

### *4.2. Evaluation Indicator*

*Accuracy*: Accuracy is the ratio of correct predictions to all legitimate samples. Accuracy indicates the performance of the authenticator in classifying legitimate users.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

where positive samples that are correctly predicted as positive are designated as True Positives (*TP*), negative samples that are correctly predicted as negative are designated as True Negatives (*TN*), negative samples that are erroneously predicted as positive are classified as False Positives (*FP*), positive samples mistakenly classified as negative are labeled as False Negatives (*FN*).

*Precision Rate*: Precision rate is an indicator of the model's capacity to accurately predict the positive samples.

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

*Recall Rate*: Recall rate measures the proportion of samples correctly identified by the model as positive classes (true classes) out of all actual positive class samples.

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

### *4.3. Performance Analysis*

Figure 5 illustrates the loss variation in the AECN and CNN across different epochs. In the initial 20 epochs, the loss of the AECN model decreases rapidly. As the number of epoch increases, the rate of loss reduction gradually slows down and ultimately trends towards 0, indicating the convergence of the AECN model. In contrast, the CNN model demonstrates a slower rate of convergence compared to the AECN model. Furthermore, within the initial 20 epochs, the CNN model exhibits larger fluctuations in loss compared to the AECN model, indicating that the AECN model exhibits superior generalization ability and is capable of demonstrating excellent performance on new data.

Figure 6 illustrates the authentication accuracy of the CNN and AECN with a authentication levels fixed at 5 and 10, respectively. At the authentication level of 5, the authentication accuracy of the CNN is 95.2%, while that of the AECN is 99.7%, representing a 4.5% increase in accuracy compared to the CNN. Upon raising the authentication level to 10, the authentication accuracy of the AECN remains at 98.3%, although this is

lower than at the authentication levels of 5. In contrast, the authentication accuracy of the CNN is 91.6%. This demonstrates that our proposed attention mechanism module is capable of maintaining excellent authentication accuracy when the security authentication level increases.

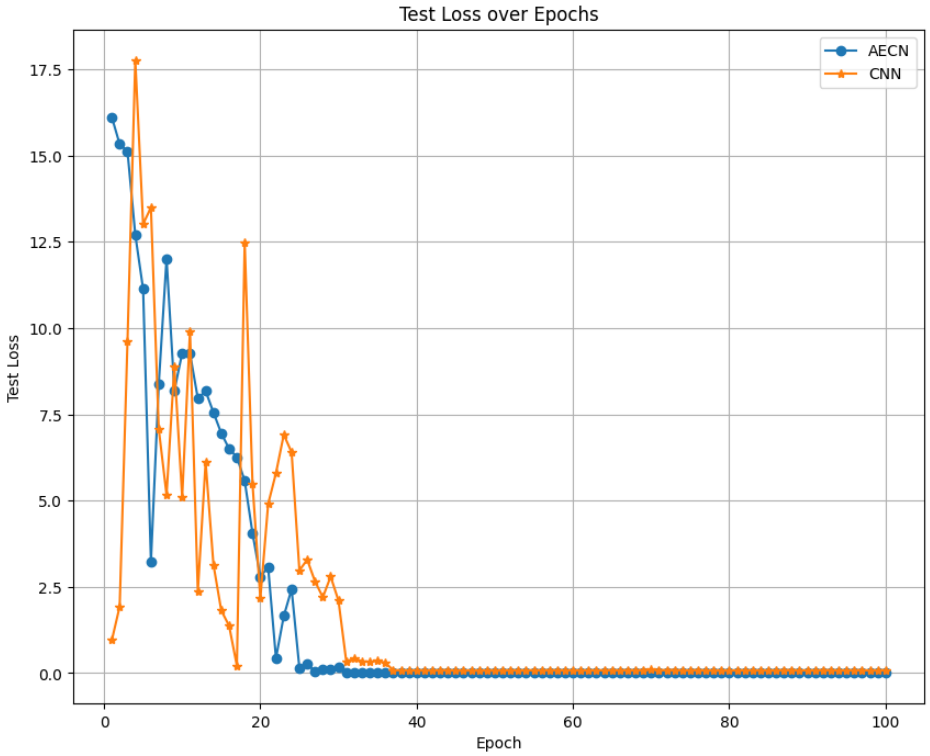


**Figure 5.** Loss of AECN and CNN.

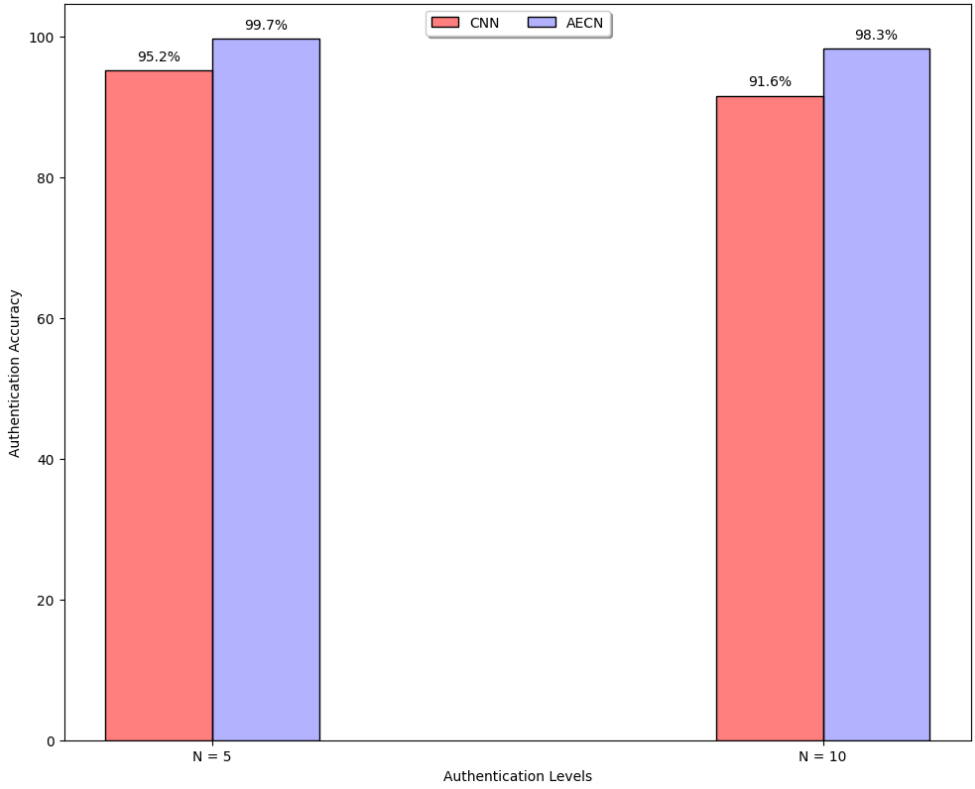


**Figure 6.** Authentication accuracy of AECN and CNN at authentication level N = 5 and N = 10.

Figure 7 shows the accuracy of the CNN and AECN when the number of authentication level is 5. In the first 20 epochs, the accuracy of the CCN fluctuated sharply, while the fluctuation of the AECN in this stage was significantly smaller than that of the CNN, indicating that the AECN is superior to the CNN in learning CSI features. After the 20th epoch, with the increase in epoch, the precision rate of both models shows an upward trend, and finally the precision rate of AECN converged to 99.7% and that of CNN converged to 92.8%.

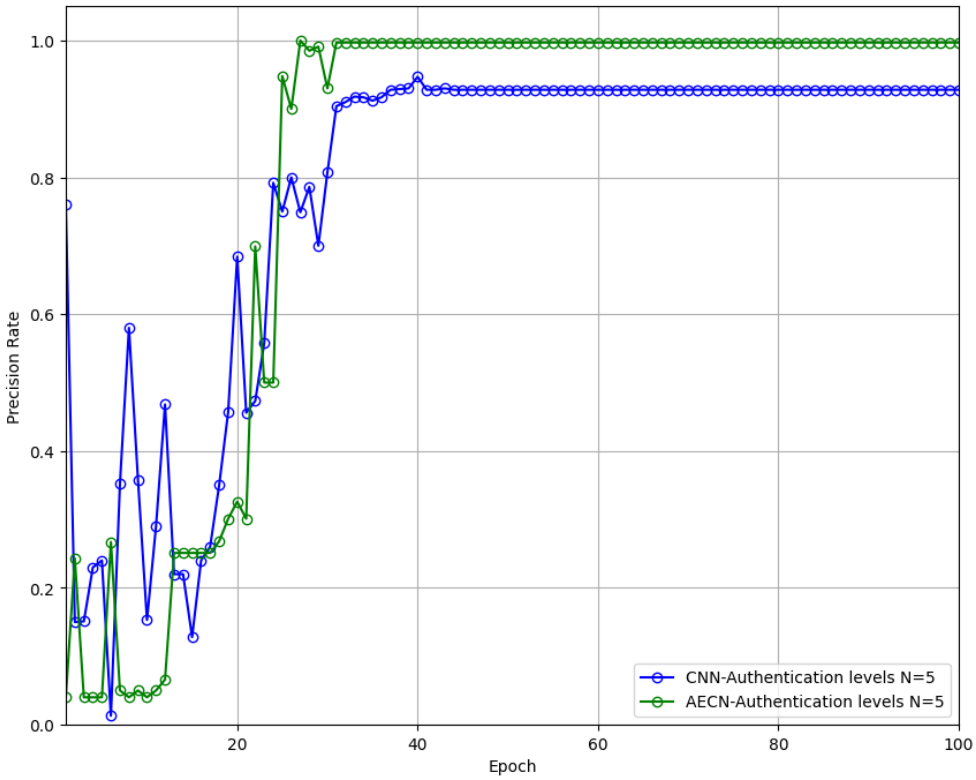


**Figure 7.** Precision rate of AECN and CNN at authentication level N = 5.

Figure 8 illustrates the recall rates of the CNN and AECN models with the authentication level fixed at 5. As the number of epochs increases, the recall rate of the AECN shows a consistent upward trend. Prior to the 20th epoch, the recall rate of the CNN exhibited significant oscillation. In terms of convergence speed, the AECN outperformed the CNN, converging more quickly. As the epochs increase, the recall rate of the AECN approaches 1, while the CNN converges at 96.3%. This indicates that our proposed AECN model demonstrates enhanced performance in distinguishing legitimate users.

Figures 9 and 10 illustrate the precision and recall rates for an authentication level of 10, respectively. Figure 9 shows that, during the initial 30 epochs, the CNN model experiences greater oscillations, ultimately converging at a precision rate of 90.4%. In comparison, the AECN model exhibits relatively minor fluctuations in its precision rate, eventually reaching a value of 99.3%. Throughout the training process, the precision rate of the AECN remains stable, showing no significant upward or downward trend. This suggests that the AECN model is better able to maintain its generalization ability.

With the authentication level set at 10, the recall rates of the CNN and AECN models are presented in Figure 10. As the number of epochs increases, the recall rates for both the AECN and CNN demonstrate a consistent upward trend. The fluctuations in the recall of the AECN are less pronounced than those of the CNN, and the AECN converges more quickly than the CNN prior to the convergence of both models. Notably, once the models have converged, the CNN achieves a recall rate of 93.2%, while the AECN attains a significantly higher recall rate of 99.3%. This indicates that our proposed AECN model,

which includes an attention module, exhibits strong performance and generalization ability, even when the security authentication level is increased to 10.

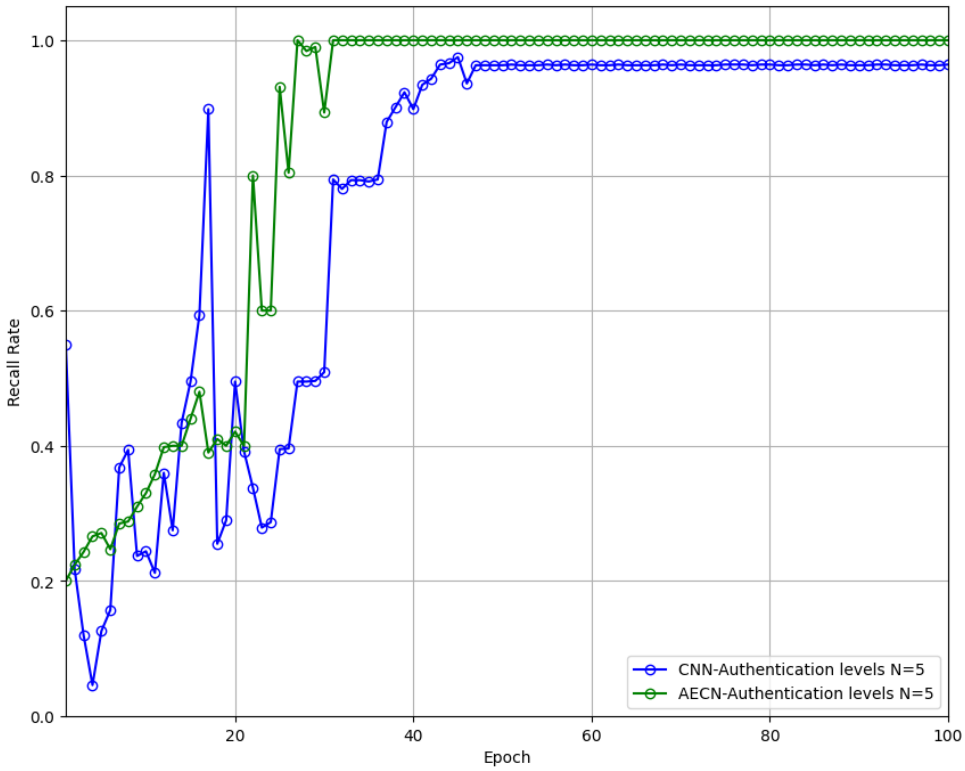


**Figure 8.** Recall rate of AECN and CNN at authentication level N = 5.

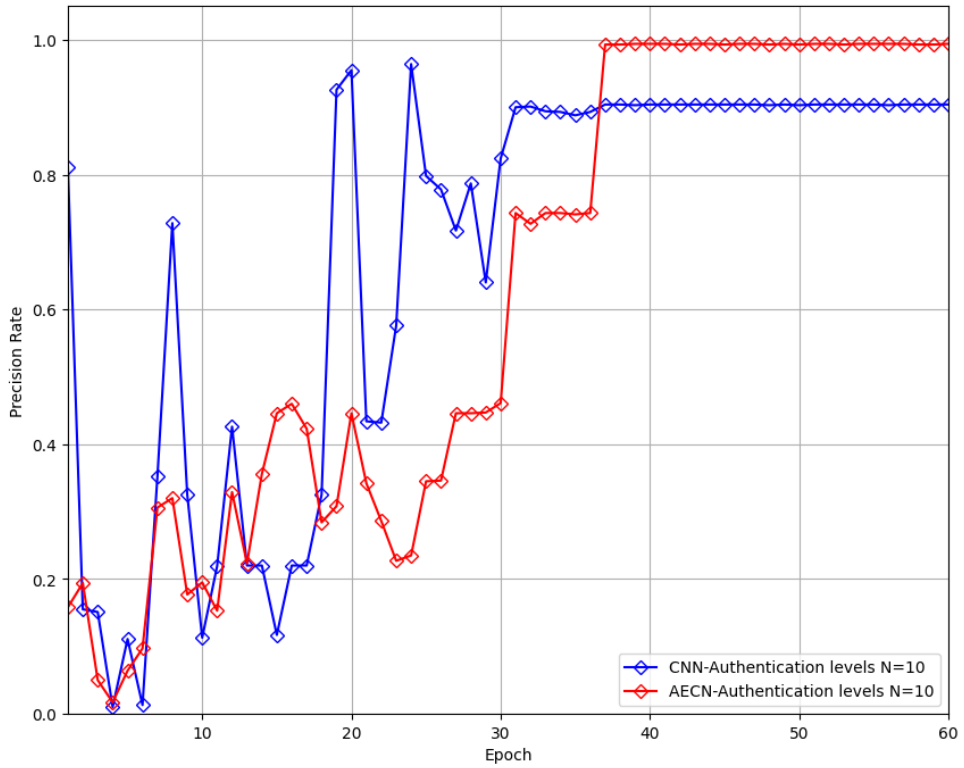


**Figure 9.** Precision rate of AECN and CNN at authentication level N = 10.

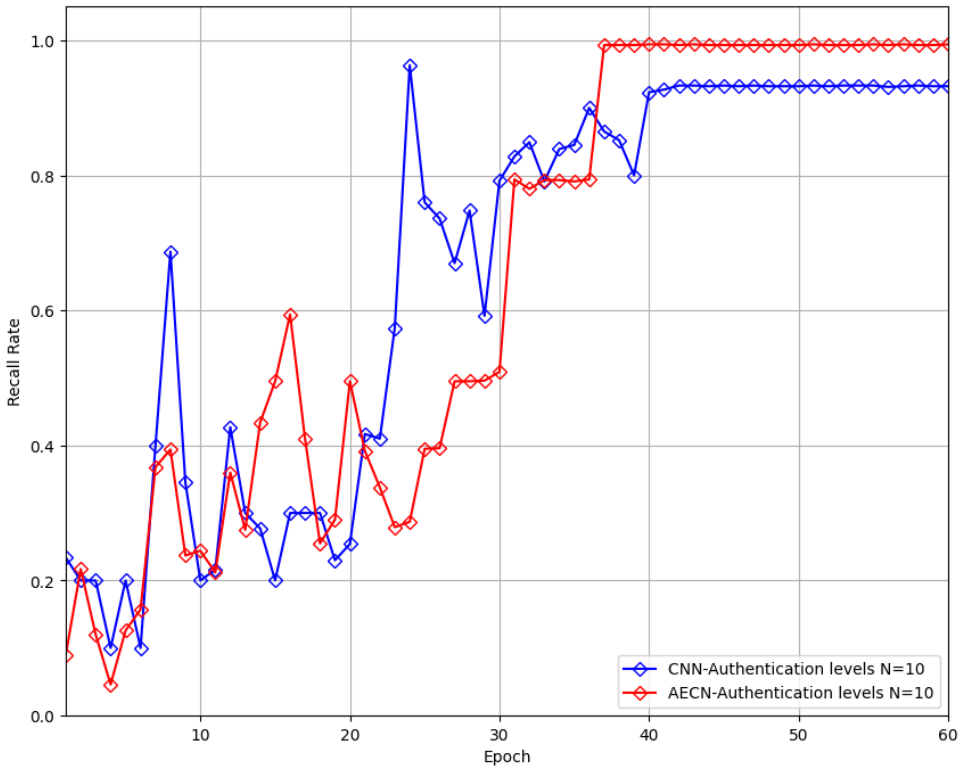


**Figure 10.** Recall rate of AECN and CNN at authentication level N = 10.

## 5. Discussion and Future Work

We designed a hierarchical security authenticator for IoT devices. Firstly, we proposed the concept of hierarchical security authentication, which enables multi-level access control by transforming the hierarchical security problem into a multi-classification problem, thereby enhancing the flexibility of security authentication for IoT devices. Secondly, the confidence score branch employs a sigmoid function to generate a confidence score between 0 and 1, indicating the degree of trust the receiver assigns to the transmitter's identity prediction. This eliminates the necessity for a fixed decision threshold in traditional authentication, thereby enhancing reliability in dynamic environments. Thirdly, an attention module comprising a bottleneck structure and a depth-separable convolution is devised for integration into the AECN model. The AECN model utilises an attention module that can capture key CSI features, generate attention weights and feature maps, and improve the classification performance of the network and the accuracy of the authentication.

However, our scheme has some aspects that require further improvement. For example, the performance of the model may be affected when dealing with different network conditions, especially in environments where noise and interference levels fluctuate. For future work, we will continue to optimize the performance of the model under different network conditions. Specifically, by utilizing data from both the simulated environment and the real environment for training, the generalization ability and robustness of the model are planned to be enhanced. In addition, we will explore methods of model optimization to reduce the computational complexity of the model without affecting its performance.

## 6. Conclusions

A hierarchical security authentication mechanism that integrates an AECN with a confidence score branch is presented. The AECN includes an attention mechanism module, which uses bottleneck structures and depth-wise separable convolution to accurately learn and focus on the crucial CSI features. This approach generates efficient feature maps and significantly enhances the network's classification performance. Additionally, we eliminate the conventional reliance on threshold detection and integrate a confidence score

component to evaluate the prediction probability. We introduce the hyperparameter λ as a tuning weight for the confidence loss to optimize the balance between the AECN's classification task and overall loss. Simulation results demonstrate that, at an authentication level of 5, the authentication accuracy of theAECN is 99.7%. When the authentication level is increased to 10, the accuracy remains high at 98.3%. Therefore, the proposed hierarchical security authentication scheme can serve as a solution for IoT security.

**Author Contributions:** X.Q.: Study Conception and Design, Writing—Review, Editing and Funding. G.Z.: Study Conception and Design, Coding, Writing—Original Draft, Writing—Review. J.Y.: Coding, Data Collection. W.J.: Writing—Review, Funding. Z.G.: Writing—Review, Funding, M.X.: Writing—Review, Funding. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the National Key R&D Program of China under grant no. 2022YFB2703000, in part by the Young Backbone Teachers Support Plan of BISTU under grant no. YBT202437, in part by the R&D Program of Beijing Municipal Education Commission under grant no. KM202211232012, and in part by the Educational Innovation Program of BISTU under grant no. 2025JGYB19.

**Data Availability Statement:** Data available upon request from the authors.

**Conflicts of Interest:** The authors declare no conflicts of interest to report regarding the present study.

## Abbreviations

| Abbreviation | Full Name |
|---|---|
| IoT | Internet of Things |
| PLA | Physical Layer Authentication |
| AECN | Attention-Enhanced Convolutional Network |
| RF | Radio Frequency |
| CFO | Carrier Frequency Offset |
| CSI | Channel State Information |
| IoV | Internet of Vehicles |
| SVM | Support Vector Machine |
| GMM | Gaussian Mixture Model |
| DNN | Deep Neural Network |
| BS | Base Station |
| OFDM | Orthogonal Frequency Division Multiplexing |

## References

1. Javed, F.; Afzal, M.K.; Sharif, M.; Kim, B.S. Internet of Things (IoT) Operating Systems Support, Networking Technologies, Applications, and Challenges: A Comparative Review. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 2062–2100. [CrossRef]
2. Amutha, B. IoT Revolutionizing Healthcare: A Survey of Smart Healthcare System Architectures. In Proceedings of the 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 1–2 November 2023; pp. 1–5. [CrossRef]
3. Eliş, B.; Çar, E.; Kasap, N.; Sarı, A.İ.; Rovshenov, A. Development of IoT-Based Smart Home Automation System with Open-Sourced Components. In Proceedings of the 2024 International Conference on Electrical, Computer and Energy Technologies (ICECET), Sydney, Australia, 25–27 July 2024; pp. 1–6. [CrossRef]
4. Chen, L.; Chen, Y.; Liang, W.; Li, X.; Li, K.C.; Wang, J.; Xiong, N. MASS: A Multi-Attribute Sketch Secure Data Sharing Scheme for IoT Wearable Medical Devices Based on Blockchain. *IEEE Internet Things J.* **2024**. [CrossRef]
5. Liu, Y.; Wang, J.; Yan, Z.; Wan, Z.; Jäntti, R. A Survey on Blockchain-Based Trust Management for Internet of Things. *IEEE Internet Things J.* **2023**, *10*, 5898–5922. [CrossRef]
6. Xie, N.; Li, Z.; Tan, H. A Survey of Physical-Layer Authentication in Wireless Communications. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 282–310. [CrossRef]
7. Zhao, Y.; Hu, Y.; Zhang, X.; Ma, Y.; Li, Y. Design and Implementation of Industrial Internet Application Layer Protocol Integrating AES and RSA Encryption Algorithm. In Proceedings of the 2023 12th International Conference of Information and Communication Technology (ICTech), Wuhan, China, 14–16 April 2023; pp. 351–356. [CrossRef]
8. Luo, H.; Li, G.; Xing, Y.; Zhang, J.; Hu, A.; Wang, X. RelativeRFF: Multi-Antenna Device Identification in Multipath Propagation Scenarios. In Proceedings of the ICC 2023—IEEE International Conference on Communications, Rome, Italy, 28 May–1 June 2023; pp. 3708–3713. [CrossRef]

9. Feng, Z.; Boxuan, S.; Wenbao, J. Review of key technology and its application of blockchain. *Chin. J. Netw. Inf. Secur.* **2018**, *4*, 22–29.
10. Niu, S.; Chen, L.; Liu, W. Attribute-Based Keyword Search Encryption Scheme with Verifiable Ciphertext via Blockchains. In Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 11–13 December 2020; Volume 9, pp. 849–853. [CrossRef]
11. Wang, T.; Wang, S.; Zhou, Z.H. Machine learning for 5G and beyond: From model-based to data-driven mobile wireless networks. *China Commun.* **2019**, *16*, 165–175. [CrossRef]
12. Lu, X.; Lei, J.; Shi, Y.; Li, W. Improved Physical Layer Authentication Scheme Based on Wireless Channel Phase. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 198–202. [CrossRef]
13. Wang, H.M.; Fu, Q.Y. Channel-Prediction-Based One-Class Mobile IoT Device Authentication. *IEEE Internet Things J.* **2022**, *9*, 7731–7745. [CrossRef]
14. Soltani, N.; Reus-Muns, G.; Salehi, B.; Dy, J.; Ioannidis, S.; Chowdhury, K. RF Fingerprinting Unmanned Aerial Vehicles with Non-Standard Transmitter Waveforms. *IEEE Trans. Veh. Technol.* **2020**, *69*, 15518–15531. [CrossRef]
15. Senigagliesi, L.; Baldi, M.; Gambi, E. Comparison of Statistical and Machine Learning Techniques for Physical Layer Authentication. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 1506–1521. [CrossRef]
16. Shi, H.; Niu, M.; Teng, Y.; Zhang, P. Privacy-Preserving Physical Layer Authentication Exploiting Carrier Frequency Offset in UAV-aided Communication Systems. In Proceedings of the 2024 International Conference on Networking and Network Applications (NaNA), Yinchuan, China, 9–12 August 2024; pp. 247–251. [CrossRef]
17. Xu, Q.; Zheng, R.; Saad, W.; Han, Z. Device Fingerprinting in Wireless Networks: Challenges and Opportunities. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 94–104. [CrossRef]
18. Xiao, L.; Greenstein, L.; Mandayam, N.; Trappe, W. Fingerprints in the Ether: Using the Physical Layer for Wireless Authentication. In Proceedings of the 2007 IEEE International Conference on Communications, New Orleans, LA, USA, 11–13 April 2007; pp. 4646–4651. [CrossRef]
19. Lu, X.; Lei, J.; Shi, Y.; Li, W. Physical-Layer Authentication Based on Channel Phase Responses for Multi-Carriers Transmission. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 1734–1748. [CrossRef]
20. Xie, N.; Zhang, Q.; Chen, J.; Tan, H. Privacy-Preserving Physical-Layer Authentication for Non-Orthogonal Multiple Access Systems. *IEEE J. Sel. Areas Commun.* **2022**, *40*, 1371–1385. [CrossRef]
21. Xiao, L.; Li, Y.; Han, G.; Liu, G.; Zhuang, W. PHY-Layer Spoofing Detection with Reinforcement Learning in Wireless Networks. *IEEE Trans. Veh. Technol.* **2016**, *65*, 10037–10047. [CrossRef]
22. Xiao, L.; Chen, T.; Han, G.; Zhuang, W.; Sun, L. Game Theoretic Study on Channel-Based Authentication in MIMO Systems. *IEEE Trans. Veh. Technol.* **2017**, *66*, 7474–7484. [CrossRef]
23. Tian, S.; Jiang, W. Blockchain Traceability Process for Hairy Crab Based on Cuckoo Filter. *Appl. Sci.* **2024**, *14*, 8027. [CrossRef]
24. Fang, H.; Zhu, Y.; Zhang, Y.; Wang, X. Decentralized Edge Collaboration for Seamless Handover Authentication in Zero-Trust IoV. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 8760–8772. [CrossRef]
25. Pei, C.; Zhang, N.; Shen, X.S.; Mark, J.W. Channel-based physical layer authentication. In Proceedings of the 2014 IEEE Global Communications Conference, Austin, TX, USA, 8–12 December 2014; pp. 4114–4119. [CrossRef]
26. Liu, H.; Wang, Y.; Liu, J.; Yang, J.; Chen, Y.; Poor, H.V. Authenticating Users Through Fine-Grained Channel Information. *IEEE Trans. Mob. Comput.* **2018**, *17*, 251–264. [CrossRef]
27. Fang, H.; Wang, X.; Xu, L. Fuzzy Learning for Multi-Dimensional Adaptive Physical Layer Authentication: A Compact and Robust Approach. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 5420–5432. [CrossRef]
28. Wang, S.; Li, N.; Xia, S.; Tao, X.; Lu, H. Collaborative Physical Layer Authentication in Internet of Things Based on Federated Learning. In Proceedings of the 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Helsinki, Finland, 13–16 September 2021; pp. 714–719. [CrossRef]
29. Wang, N.; Jiang, T.; Lv, S.; Xiao, L. Physical-Layer Authentication Based on Extreme Learning Machine. *IEEE Commun. Lett.* **2017**, *21*, 1557–1560. [CrossRef]
30. Pan, F.; Pang, Z.; Wen, H.; Luvisotto, M.; Xiao, M.; Liao, R.F.; Chen, J. Threshold-Free Physical Layer Authentication Based on Machine Learning for Industrial Wireless CPS. *IEEE Trans. Ind. Inform.* **2019**, *15*, 6481–6491. [CrossRef]
31. Qiu, X.; Jiang, T.; Wu, S.; Hayes, M. Physical Layer Authentication Enhancement Using a Gaussian Mixture Model. *IEEE Access* **2018**, *6*, 53583–53592. [CrossRef]
32. Xiao, L.; Wan, X.; Han, Z. PHY-Layer Authentication with Multiple Landmarks with Reduced Overhead. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 1676–1687. [CrossRef]
33. Meng, R.; Xu, X.; Sun, H.; Zhao, H.; Wang, B.; Han, S.; Zhang, P. Multiuser Physical-Layer Authentication Based on Latent Perturbed Neural Networks for Industrial Internet of Things. *IEEE Internet Things J.* **2023**, *10*, 637–652. [CrossRef]
34. 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.