On Physical-Layer Authentication via Online Transfer Learning

Yi Chen, *Student Member, IEEE*, Pin-Han Ho, *Fellow, IEEE*, Hong Wen, *Senior Member, IEEE*, Shih Yu Chang, *Senior Member, IEEE*, and Shahriar Real

***Abstract*—This article introduces a novel physical-layer (PHY-layer) authentication scheme, called transfer learning-based PHY-layer authentication (TL-PHA), aiming to achieve fast online user authentication that is highly desired for latency-sensitive applications such as edge computing. The proposed TL-PHA scheme is characterized by incorporating with a novel convolutional neural network architecture, namely, the triple-pool network (TP-Net), for achieving lightweight and online classification, as well as effective data augmentation methods for generation of data set samples for the network model training. To assess the performance of the proposed scheme, we conducted two sets of experiments, including the one using computer-simulated channel data and the other utilizing real experiment data generated by our wireless testbed. The results demonstrate the superiority of the proposed scheme in terms of authentication accuracy, detection rate, and training complexity compared to all the considered counterparts.**

***Index Terms*—Channel state information (CSI), convolutional neural network (CNN), edge computing (EC), physical-layer (PHY-layer) authentication, transfer learning (TL).**

I. Introduction

Edge computing (EC) is an emerging technology aiming to extend the legacy cloud computing services to the network edge and resolve some of its inherent limitations, such as real-time control incompetence, network traffic bottlenecks, and cloud data privacy insecurity [1]–[3]. An EC system is generally formed by a set of edge devices and Internet of Things (IoT) terminal units, targeting at achieving a low-latency and high-reliability service provisioning platform for supporting heterogeneous EC applications, such as smart grid, smart home, smart healthcare, and vehicle networks [3]–[5]. Some edge devices of high capacity may serve as small/micro data centers that are equipped with the intelligence to ensure Quality of Service (QoS), while the IoT terminals are generally with limited resources in terms of computational power and memory that could be subject to various security threats in the presence of the openness of wireless media.

Physical-layer (PHY-layer) authentication has been considered an effective alternative to the conventional key-based approaches by exploring the rich PHY-layer channel characteristics between the EC and client device (CD) [2]. Numerous researches on PHY-layer authentication have been reported [6]–[10], mostly focusing on the manipulation of novel statistical models/algorithms of binary hypothesis tests, whose accuracy highly relies on the test thresholds.

Alternatively, machine learning (ML)-based PHY-layer authentication has been studied, which aims to achieve high accuracy of message authentication via a well-trained ML model [11]–[14]. Although promising, most previously reported ML-based PHY-layer authentication schemes have employed ML models of a large number of layers and parameters, leading to significant complexity and power consumption. The size of the ML model serves as a key to the success of the considered application scenario, where a fast and adaptive online classification process is essential.

As a remedy, there have been some advanced convolutional neural network (CNN) architecture and its activation functions, such as the global-connected net (GC-Net) [15], reported to achieve much improved efficiency in both training and classification stages. Furthermore, sufficient data sets are essential to obtain a well-trained model, which are nonetheless not always available in the highly dynamic EC environment. As such, efficient data augmentation schemes that encompass a suite of techniques, such as geometric transformations, kernel filters, random erasing, feature space augmentation, and adversarial training, are highly desired [16]–[19]. With more training, data sets definitely lead to longer training time. In order to fit into the real-time scenario, we consider transfer learning (TL) for the proposed EC system, where the EC is allowed to train the model in off-line and the training result can be migrated to the network model in real-time. This article introduces a novel TL-based PHY-layer scheme, called, TL-PHA, for lightweight user/message authentication. It is uniquely featured by a number of novel designs, aiming to resolve all the above-mentioned issues, which are summarized as follows.

Manuscript received March 22, 2021; revised May 12, 2021; accepted May 26, 2021. Date of publication June 4, 2021; date of current version January 7, 2022. This work was supported in part by China Scholarship Council and in part by the National Major Research and Development Program, China, under Grant 2018YFB0904900 and Grant 2018YFB0904905. This article was presented in part at the 2020 IEEE Globecom Workshops (GC Wkshps), Taipei, Taiwan, December 2020. *(Corresponding author: Hong Wen.)*

Yi Chen is with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China, and also with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: chenyi1309@126.com).

Pin-Han Ho and Shahriar Real are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: p4ho@uwaterloo.ca; shahriar.real@uwaterloo.ca).

Hong Wen is with the School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: sunlike@uestc.edu.cn).

Shih Yu Chang is with the Department of Applied Data Science, San Jose State University, San Jose, CA 95192 USA (e-mail: shihyu.chang@sjsu.edu).

Digital Object Identifier 10.1109/JIOT.2021.3086581

1) We claim the proposed TL-PHA is the first PHY-layer authentication scheme for the EC applications that employs the TL strategy. Distinguished from the existing TL strategies, the proposed TL-PHA scheme can determine the moment that online fine tuning for the network model is needed according to the offline training results.
2) To enable a fast training and classification process, we introduce a novel CNN interconnection architecture, called TP-Net, to facilitate high-efficiency model training and online classification.
3) To make up the training data sets insufficiency, two data augmentation methods are investigated and incorporated in the proposed TL-PHA scheme.
4) Extensive simulation is conducted to verify the proposed TL-PHA scheme via both computer simulation and USRP testbed, respectively.

The remainder of this article is organized as follows. Section II reviews the related studies of PHY-layer authentication, TL, data augmentation, and CNN. Section III illustrates the system model of the study. Section IV presents the proposed TL-PHA scheme. Section V shows the introduction of the triple pool network (TP-Net). Section VI introduces the data augmentation methods employed in the proposed TL-PHA. Section VII shows the experiment results and comparison with the counterparts. This article is concluded in Section VIII.

II. Related Work

A. PYH-Layer Authentication

A number of PHY-layer authentication schemes were reported in [6]–[12]. The research approach of the binary hypothesis test upon a given threshold is relatively mature. A representative of this research direction is given in [6]–[8], where PHY-layer authentication is performed via exploiting the spatial variability of the radio channel between unknown channel state information (CSI) and known legal CSI. Xiao *et al.* [6] proposed a generalized likelihood ratio test version for PHY-layer authentication. Wen *et al.* [7] introduced two cross-layer authentication schemes for the smart meter system, namely, the symmetric cryptography-based PHY-layer-assisted authentication scheme and the public-key infrastructure-based PHY-layer-assisted authentication scheme. Pan *et al.* [8] applied PHY-layer authentication based on CSI to measurements from indoor, outdoor, moving, and stationary industrial wireless communication scenarios, respectively. They derived some meaningful insights on the applicability of the PHY-layer method to industrial wireless communications via the CSI analysis that still depended on the threshold. Nevertheless, the authentication accuracy of these approaches is highly rely on the test threshold values that are hard to obtain in practical environment.

Adaptive classification approaches based on ML are gaining popularity as new strategies for the PHY-layer authentication scenarios. Pan *et al.* [9] proposed a threshold-free PHY-layer authentication scheme based on ML for the industrial mobile scenario, which can replace the traditional threshold-based decision-making method. An adaptive PHY-layer authentication scheme based on ML as an intelligent process to learn and use the complex time-varying environment was proposed in [10]. A spoofing detection scheme based on reinforcement learning process was introduced in [11]. Wang *et al.* [12] utilized a deep neural network to complete the indoor location via CSI. Liao *et al.* [2] proposed a data-augmented multiuser PHY-layer authentication scheme based on the deep neural network. The above schemes, although being claimed effective in the considered scenarios, used conventional ML models with a large number of layers and parameters. For an ML model to be applied under latency-sensitive applications such as EC, the efficiency of such an ML model cannot be just measured by ML metric, we should also consider the required resources to perform model training and prediction simultaneously.

B. Transfer Learning

TL has been reported to resolve the situation by using what is learned for one problem to assist another different but related problem. It also has been widely applied to the image processing scenario [20]–[22]. The main two ways of doing TL in deep neural networks are the fine-tuning method and the freezing layer method, respectively. It requires to retrain the whole network parameters in the fine-tuning method. Instead, the freezing-layer method freezes most of the transferred parameters [23].

The conventional ML algorithms need to be trained from scratch every time to solve specific tasks. However, training a neural network from scratch may be cumbersome and the available data sets may not be rich enough to effectively capture model features. Therefore, the trained neural model could not be generalized properly when applied in the practical contexts.

Even if there are enough data sets to train the network online in real time, it is also time consuming to train the model by the server, which is not proper for latency-sensitive applications. Therefore, in this article, we exploit data sets to train the ML model offline, and utilize the trained model parameters to perform a fast and accurate online PHY-layer authentication. This is our core idea to apply TL in PHY-layer authentication.

C. Data Augmentation

Data augmentation is the technique of increasing the size of data set used for training a deep learning model [16]–[19]. For desired predictions, the ML models often require sufficient training data set, which is not always available. Therefore, the existing data set is augmented in order to make a better generalized model. Frid-Adar *et al.* [19] utilized GAN-based image synthesis data augmentation to improve classification performance for liver lesion classification in 2018. Li *et al.* [17] exploited data augmentation approaches to enhance the authentication accuracy on smartphones. The audio data augmentation for overcoming the problem of data scarcity on environmental sound classification was introduced in [18]. However, these conventional data augmentation methods mainly include flip, rotation, scale, kernel filters, random erasing, and so on, which are not proper to increase the diversity of PHY-layer CSI, because they will destroy the correlation structure between the channel matrices.

D. Convolutional Neural Network

CNN has been well recognized as a powerful tool for intelligent decision making in the presence of complex data. Krizhevsky *et al.* [24] first applied a CNN to achieve great success on ImageNet competition in 2012. Simonyan and Zisserman [25] found the deep convolutional network (i.e., VGG) has a significant effect on the increase in accuracy for image recognition tasks. A residual-network architecture (ResNet) was proposed to alleviate the training time of networks in [26]. A deep CNN architecture code named Inception [27] was proposed for the reduction of computational burden of the network architecture. Nonetheless, most these reported CNN architectures still require numerous layers and a huge number of parameters to achieve the desired accuracy, and are thus subject to extremely high computation complexity.

GC-Net [15] is a recently introduced CNN architecture that is reported to achieve similar performance to that by its predecessors while taking much less parameters and, thus, much less computation resources. The unique features of GC-Net include a globally connected interconnection architecture and a piecewise-linear activation function between the convolution layers that can successfully mitigate the gradient-vanishing problem. However, it still has room for improvement, especially in the application of PHY-layer authentication.

III. System Model

We consider the scenario of latency-sensitive EC systems, where an edge computing node (ECN) is associated with multiple CDs. Authentication is needed when any message is delivered between the ECN and each CD to avoid any malicious attempt to the EC system.

The key element of PHY-layer authentication is the CSI (i.e., channel response matrix H) obtained via a channel estimation procedure [28]–[30]. The signal of the receiver is given by

$$y_p(t) = h_p \times x_p(t) + n(t) \tag{1}$$

where y_p is the received signal at the ECN, t is the time interval between every data frames, h_p refers to a time-domain channel matrix containing the channel coefficients, x_p represents the pilot signal known to both CDs and ECN used to estimate the CSI, and $n(t)$ is the additive white Gaussian noise with variance σ^2. The channel time-domain response estimated by the ECN through the channel estimation is as follows:

$$\begin{aligned} h(t) &= y_p(t) \times x_p^{-1}(t) \\ &= h_p \times x_p(t) \times x_p^{-1}(t) + n(t) \times x_p^{-1}(t) \\ &= h_p + n(t) \times x_p^{-1}(t) \end{aligned} \tag{2}$$

where $x_p^{-1}(t)$ is the inversion of $x_p(t)$. The ECN is able to get the channel frequency response matrix, i.e., CSI H, by the discrete Fourier transform (DFT) of the time-domain signal.

For the sake of simplicity, the channel frequency response matrix, i.e., CSI H, is denoted as

$$H = Y_p X_p^{-1} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{bmatrix} \tag{3}$$

where X_p is the known pilot and Y_p is the pilot obtained at the ECN. By assuming orthogonal frequency-division multiplexing (OFDM), where $H \in \mathbb{C}^{m\times n}$, $z_{uv} = a_{uv} + jb_{uv}$, $a, b \in \mathbb{R}$, $u = 1, 2, \ldots, m$, $v = 1, 2, \ldots, n$, $\mathrm{j}^2 = -1$, and $m = N_s \times N_t$, with N_s being the number of subcarriers and N_t being the number of transmitting antennas, whereas n is the number of receiving antennas.

The ECN first validates the received signal Y_p via the conventional cryptography-based approach implemented by the upper layer, and goes through the CSI extraction. Once prepared, the PHY-layer authentication can be performed on the incoming messages. The detailed process of upper level authentication is given as follows.

1) In the beginning, the timer is set to 0, i.e., $t = 0$. The flag bit representing the completion of PHY-layer authentication preparation is also set to 0, i.e., $F_{\mathrm{PHY}} = 0$, where F_{PHY} is set to 1 when the PHY-layer authentication is ready to perform. In addition, the counter of the data frame is set to 0, i.e., $\mathrm{num}_{\mathrm{data}} = 0$.
2) Upon receiving a message from a CD, the ECN first checks the flag bit of F_{PHY}, where the proposed PHY-layer authentication is performed if $F_{\mathrm{PHY}} = 1$, and the upper level numerical authentication is relegated otherwise. In the former, the ECN has to extract the CSI H_k^t, the channel response vector of the kth node at time t.
3) If the upper level authentication is not successful, the message is considered to be illegal and discarded; otherwise, the ECN considers the message legal, measures, and records the CSI H_k^t.

The collected channel response matrices shall be used to train the TP-Net model offline. The process of upper level authentication is summarized in Algorithm 1.

IV. Proposed TL-PHA

The proposed PHY-layer authentication scheme incorporates with three functional modules as shown in Fig. 1: 1) offline training; 2) online migration and fine tuning; and 3) online decision making. Before offline training, ECN needs to collect CSIs, whose legitimacy is determined via upper level authentication (i.e., running Algorithm 1).

A. Offline Training

In the offline training phase, the ECN first uses the channel frequency response matrices H to generate the channel training vectors specific to each CD, which are, in turn, used for training the TP-Net model. The details are as follows.

1) The ECN first adjusts the complex matrix H to get a real matrix U with a size of $m \times 2n$

$$U = [P, Q] = [\Re(H), \Im(H)] \tag{4}$$

Algorithm 1 Upper Level Authentication

```
Input: Wireless signal carrying information sent from CD.
Output: Channel frequency response matrix, i.e., CSI $H$.
Initialization parameters.
Upon receiving a message from a CD, ECN checks the PHY-layer authentication flag bit:
if $F_{PHY} = 1$ then
    ECN extracts the CSI $H_k^t$ and performs the PHY-layer authentication with the well trained TP-Net model, i.e., running Algorithm 4.
else if $F_{PHY} = 0$ then
    ECN activates the upper level authentication.
    if The authentication is unsuccessful then
        The message is illegal and discarded.
    else
        ECN considers the message legal, measures and records the CSI $H_k^t$.
    end if
end if
Output channel frequency response matrices.
```

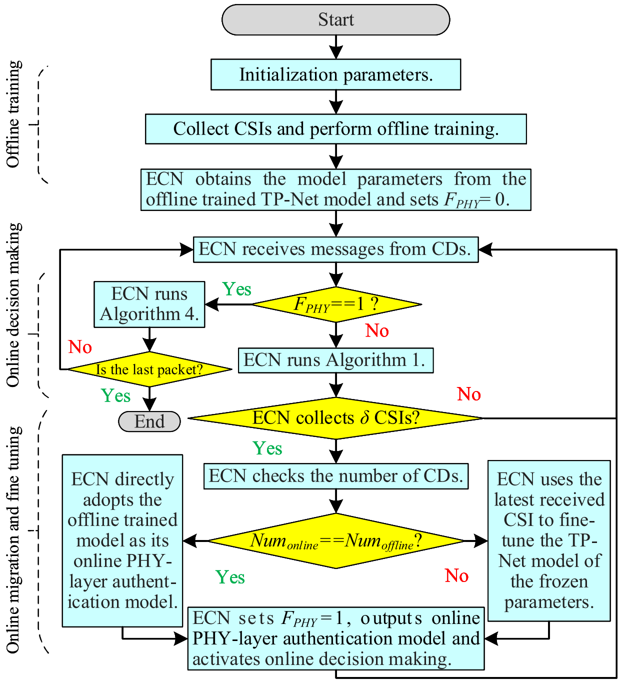

Start
Initialization parameters.
Collect CSIs and perform offline training.
ECN obtains the model parameters from the offline trained TP-Net model and sets F_{PHY}= 0.
ECN receives messages from CDs.
F_{PHY}==1 ?
Yes
ECN runs Algorithm 4.
No
ECN runs Algorithm 1.
No
Is the last packet?
Yes
End
ECN collects δ CSIs?
No
Yes
ECN checks the number of CDs.
Num_{online}==$Num_{offline}$?
Yes
No
ECN directly adopts the offline trained model as its online PHY-layer authentication model.
ECN uses the latest received CSI to fine-tune the TP-Net model of the frozen parameters.
ECN sets F_{PHY}=1, outputs online PHY-layer authentication model and activates online decision making.
Offline training
Online decision making
Online migration and fine tuning

Fig. 1. Flowchart of TL-PHA.

where $\Re(\cdot)$ represents the real part, $\Im(\cdot)$ is the imaginary part, $P = \Re(H)$, $Q = \Im(H)$, so the formula of U is as follows:

$$U = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1n} \\ a_{21} & \cdots & a_{2n} & b_{21} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mn} \end{bmatrix}. \quad (5)$$

The superscript and subscript of U_k^t are specific to the kth node at time t of the channel response, respectively.

Algorithm 2 Offline Training for PHY-Layer Authentication

```
Input: Channel frequency response matrices $H$.
Output: Authentication model.
ECN adjusts the complex channel matrix $H_k^t$ to get a real channel matrix $U_k^t$.
ECN gives the new real channel matrix $U_k^t$ a label $I_k$.
ECN generates training set $\boldsymbol{D}_{train}$.
ECN employs $\boldsymbol{D}_{train}$ to train the TP-Net model.
Output authentication model.
```

2) A label I_k is given to the newly obtained real channel matrix U_k^t, which is a unit vector represented by an one-hot code.

3) The training set $\boldsymbol{D}_{\text{train}}$ is generated as follows:

$$\boldsymbol{D}_{\text{train}} = \{\boldsymbol{X}_{\text{train}}, \boldsymbol{Y}_{\text{train}}\} \quad (6)$$

where

$$\boldsymbol{X}_{\text{train}} = \begin{bmatrix} \mathbf{U}_1^{t_1} \\ \mathbf{U}_2^{t_2} \\ \vdots \\ \mathbf{U}_k^{t_k} \end{bmatrix} = \begin{bmatrix} U_1^1 & U_1^2 & \cdots & U_1^{\psi_1} \\ U_2^1 & U_2^2 & \cdots & U_2^{\psi_2} \\ \vdots & \vdots & \ddots & \vdots \\ U_k^1 & U_k^2 & \cdots & U_k^{\psi_k} \end{bmatrix} \quad (7)$$

$$\boldsymbol{Y}_{\text{train}} = \begin{bmatrix} \mathbf{I}_1^{t_1} \\ \mathbf{I}_2^{t_2} \\ \vdots \\ \mathbf{I}_k^{t_k} \end{bmatrix} = \begin{bmatrix} I_1^1 & I_1^2 & \cdots & I_1^{\psi_1} \\ I_2^1 & I_2^2 & \cdots & I_2^{\psi_2} \\ \vdots & \vdots & \ddots & \vdots \\ I_k^1 & I_k^2 & \cdots & I_k^{\psi_k} \end{bmatrix}. \quad (8)$$

4) The training set $\boldsymbol{D}_{\text{train}}$ is used to train the TP-Net model that will be used for the future online processes.

The offline training for PHY-layer authentication is summarized in Algorithm 2.

B. Online Migration and Fine Tuning

The ECN periodically checks if the number of CDs is changed. If not, the ECN does not need to update the trained model parameters. Otherwise, the ECN uses some of the most updated CSI to fine tune the TP-Net model based on the offline training result. The detailed process of the online migration and fine tuning is given as follows:

1) ECN first obtains the model parameters from the offline-trained TP-Net model and resets the flag bit $F_{\text{PHY}} = 0$.

2) ECN online receives messages and runs Algorithm 1 for collecting CSIs. If ECN collects δ (e.g., $\delta \geq 3$) CSIs per CD, it turns to the step of checking the number of CDs; otherwise, waits for the next new message.

3) If $\text{Num}_{\text{online}} = \text{Num}_{\text{offline}}$, ECN directly adopts the offline trained model as its online PHY-layer authentication model and sets $F_{\text{PHY}} = 1$, representing the readiness of PHY-layer authentication for the subsequent incoming messages.

4) Otherwise, the ECN uses the collected CSIs to fine tune the model parameters in the last layer in the offline trained model while freezing all the others.

Algorithm 3 Online Migration and Fine Tuning for PHY-Layer Authentication

```
Input: The offline well trained TP-Net model.
Output: Online authentication model.
ECN obtains model parameters from the offline trained TP-Net model.
ECN online receives message and performs Algorithm 1.
ECN checks the number of CSIs:
if ECN collects $\delta$ (e.g., $\delta \geqslant 3$) CSIs per CD. then
    ECN checks the number of CDs:
    if Num$_{online}$ = Num$_{offline}$ then
        Adopt the offline trained model as its online PHY-layer authentication model and set $F_{PHY} = 1$.
    else
        ECN utilizes the collected CSIs to fine-tune the model of frozen parameters.
        ECN takes the new trained model as its online PHY-layer authentication model, activates the online decision making and sets $F_{PHY} = 1$.
    end if
else
    ECN waits for next new message.
    if The number of CDs is changed. then
        ECN sets $F_{PHY} = 0$.
    end if
end if
Output online authentication model.
```

5) When the fine tuning is completed, the ECN activates the online decision making and sets $F_{\text{PHY}} = 1$, where the desired PHY-layer authentication can be performed on the subsequent messages.

The proposed online migration and fine-tuning mechanism is summarized in Algorithm 3.

C. Online Decision Making

With $F_{\text{PHY}} = 1$, the proposed PHY-layer authentication can start and is described as follows.

1) ECN starts to authenticate a new message via the most updated model. ECN first extracts the channel response matrix H^t_{unknown} via performing Algorithm 1.
2) The ECN adjusts the channel response matrix H^t_{unknown} as the real number matrix U^t_{unknown} like formula (5).
3) The real number channel matrix is fed into the online fine-tuned TP-Net for authentication, where the input of TP-Net is marked as $\boldsymbol{X}_{\text{Auc}}$

$$\boldsymbol{X}_{\text{Auc}} = \left[U^1_{\text{unknown}}, U^2_{\text{unknown}}, U^3_{\text{unknown}}, \ldots,\right]. \tag{9}$$

4) Finally, the ECN makes the authentication decision based on the output of $\boldsymbol{X}_{\text{Auc}}$ of the TP-Net, where the output is marked as $\hat{\boldsymbol{Y}}_{\text{Auc}}$, but the real label of the output is $\boldsymbol{Y}_{\text{Auc}}$.

The online decision making for PHY-layer authentication is summarized in Algorithm 4.

Algorithm 4 Online Decision Making for PHY-Layer Authentication

```
Input: Channel matrices of unknown nodes.
Output: Labels of unknown nodes.
for ECN receives messages sent from unknown CD do
    ECN checks the PHY-layer authentication flag bit:
    if $F_{PHY} = 1$ then
        ECN estimates the CSI $H^t_{unknown}$;
        ECN adjusts $H^t_{unknown}$ as $U^t_{unknown}$ like (5).
        The real number channel matrix of unknown node $\boldsymbol{X}_{\text{Auc}}$ is fed into the TP-Net for authentication.
        Output the label of unknown node.
    else if $F_{PHY} = 0$ then
        Perform Algorithm 1.
    end if
end for
```

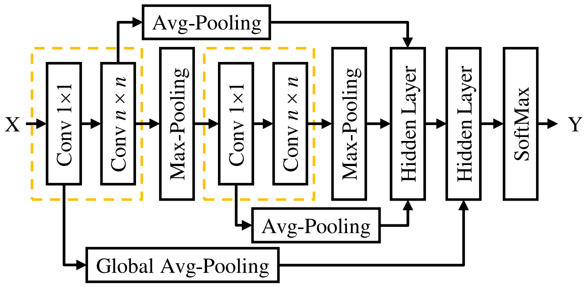


Fig. 2. Architecture of TP-Net.

V. Introduction of TP-Net

Inspired by the GC-Net [15], we propose a new CNN architecture, namely, TP-Net, whose interconnection architecture is shown in Fig. 2. With n convolutional blocks in total, each block has two filters, followed by batch normalization and activation, where any convolutional block is allowed to connect with a hidden layer that feeds into the last hidden layer and output (SoftMax) layer.

The proposed TP-Net employs an exponential linear unit (ELU) and a standard unit softmax function (SoftMax) as its activation functions, where the ELU is defined as a nonlinear function as presented in formula (10), and the SoftMax is defined by the formula (11), where $i = 1, 2, \ldots, N$, and $\boldsymbol{w} = (w_1, w_2, \ldots, w_N) \in \mathbb{R}^N$

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha(e^x - 1), & \text{if } x < 0 \end{cases} \tag{10}$$

$$\text{SoftMax}(\mathbf{w})_i = \frac{e^{w_i}}{\sum_{j=1}^{N} e^w_j}. \tag{11}$$

Distinguished from any previously reported CNN architecture, the proposed TP-Net is uniquely featured by consisting of three types of pooling: 1) max; 2) average; and 3) global. The max pooling is applied between two convolutional blocks to reduce the feature map sizes. The average pooling aims to tune the variance in the data set for possibly dimensional reduction. The global average pooling (GAP) decreases the dimension rapidly to possibly save computation cost. The main reason behind using the triple pools is that these three types of methods can utilize their strength parts to mitigate their weakness parts mutually. Thus, TP-Net model can work on the CSI database.

VI. Proposed Data Augmentation

In view of the application scenarios of edge servers with different resources, two novel data augmentation methods, namely, stochastic weight data augmentation (SWDA) and block swap data augmentation (BSDA), are introduced for the exploration of some unique features of the PHY-layer CSI, such as its time continuity and correlation between subcarriers.

A. SWDA

To explore the correlation among the CSIs of consecutive frames, SWDA generates new channel matrices using the channel response vectors of two or more consecutive frames. Stochastic weight averaging data augmentation (SWA-DA) [2] is a recently proposed data augmentation method for multiuser PHY-layer authentication. Different from SWA-DA that requires all the original data sets (CSIs) to be within a common coherence time [31], the proposed SWDA has adjacent CSIs to be within common coherence time, which not only leads to easier implementation, but can also sufficiently exploit the correlation of CSIs in adjacent frames. The detailed SWDA is given as follows.

Let $\boldsymbol{D}_k$ denote the original data set sample of the kth CD

$$\boldsymbol{D}_k = \{\boldsymbol{X}_k, \boldsymbol{Y}_k\} \tag{12}$$

where $\boldsymbol{X}_k$ is the original input samples

$$\boldsymbol{X}_k = \mathbf{U}_k^t = \left\{U_k^1, \quad U_k^2, \quad \ldots, \quad U_k^{\psi_k}\right\} \tag{13}$$

and $\boldsymbol{Y}_k$ is the original output samples

$$\boldsymbol{Y}_k = \mathbf{I}_k^t = \left\{I_k^1, I_k^2, \ldots, I_k^{\psi_k}\right\}. \tag{14}$$

The new channel response matrix is constructed according to the following formula:

$$S_k^{\mu} = \sum_{i=\mu}^{\mu+\theta-1} \left(\lambda_{i-\mu+1} \times U_k^i\right), 1 < \theta \leq \psi_k \tag{15}$$

where θ, which is a positive integer and $1 < \theta \leq \psi_k$, represents the number of original adjacent samples involved in generating a new sample, ψ_k is the total number of original samples of the kth node, λ is stochastic weight and $\sum_{i=1}^{\theta} \lambda_i = 1$, and μ is the index of new samples, natural number and $\mu = 1, 2, \ldots, \psi_k - \theta + 1$. After reconstructing the channel response matrix, the new input sample is

$$\begin{aligned} \boldsymbol{X}_{\mathrm{SWDA\_}k} &= \left\{\mathbf{U}_k^t, \mathbf{S}_k^{\mu}\right\} \\ &= \left\{U_k^1, U_k^2, \ldots, U_k^{\psi_k}, S_k^1, S_k^2, \ldots, S_k^{\mu_k}\right\}. \end{aligned} \tag{16}$$

The new output sample is

$$\boldsymbol{Y}_{SWDA\_k} = \left\{I_k^1, I_k^2, \ldots, I_k^{\psi_k}, I_k^1, I_k^2, \ldots, I_k^{\mu_k}\right\} \tag{17}$$

where $\mu_k = \psi_k - \theta + 1$.

```
Algorithm 5 PHY-Layer Authentication With SWDA
Input: The channel matrices of unknown nodes $\boldsymbol{X}_{Auc}$.
Output: The labels of unkonwn nodes $\hat{\boldsymbol{Y}}_{Auc}$.
Get the original CSIs samples $\boldsymbol{D}_{train} = \{\boldsymbol{X}_{train}, \boldsymbol{Y}_{train}\}$ and $\theta$.
ECN reconstructs the new training data set $\boldsymbol{D}_{SWDA\_train}$ according to equation (15), (16) and (17).
ECN uses the new training data set $\boldsymbol{D}_{SWDA\_train}$ by executing the functional module of offline training as given in Section IV-A.
Obtain $\boldsymbol{X}_{Auc}$.
$\boldsymbol{X}_{Auc}$ is fed into the authentication model.
Output the labels of unkonwn nodes $\hat{\boldsymbol{Y}}_{Auc}$.
```

The new training data set of the kth CD can be obtained as follows:

$$\boldsymbol{D}_{SWDA\_k} = \left\{\boldsymbol{X}_{SWDA\_k}, \boldsymbol{Y}_{SWDA\_k}\right\}. \tag{18}$$

The PHY-layer authentication with SWDA is summarized in Algorithm 5.

B. BSDA

The conventional data augmentation techniques, such as flip, rotation, and scale [16]–[18], may cause the receiver unable to correctly demodulate/decode according to the new channel matrix since the correlation structure of the channel matrix is destroyed. Motivated by the observation, BSDA is designed to generate new channel vectors with the prior and post adjacent CSIs, where some elements of two or more channel response vectors are swapped to obtain new channel response matrices. Compared to the conventional data augmentation, the BSDA scheme does not change the correlation between channel matrices, but only exchanges elements between the same position of the channel matrices, which can generate more new data samples that facilitate boosting the authentication rate. The detailed BSDA is shown as follows.

Let $\boldsymbol{D}_k$ denote the original training sample of the kth CD, and U_k^t with a size of $m \times 2n$ denote the original real number channel matrix of kth CD at t time slot and $t = 1, 2, \ldots, \psi_k$.

First, U_k^t is divided into two blocks equally according to row vectors

$$U_k^t = \begin{bmatrix} \underbrace{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_{\frac{m}{2}}}_{1}, \underbrace{\boldsymbol{c}_{\frac{m}{2}+1}, \ldots, \boldsymbol{c}_m}_{2} \end{bmatrix}^T = \begin{bmatrix} \boldsymbol{r}_k^{(t)(1)} \\ \boldsymbol{r}_k^{(t)(2)} \end{bmatrix} \tag{19}$$

where $\boldsymbol{r}_k^{(t)(1)}$ with a size of $(m/2) \times 2n$ represents the upper half block of matrix U_k^t and $\boldsymbol{r}_k^{(t)(2)}$ with size of $(m/2) \times 2n$ is the lower half block of matrix U_k^t.

Second, the partial block elements of each pair of channel matrices are swapped in adjacent η $(2 \leq \eta \leq \psi_k)$ channel matrices according to the following formula:

$$\begin{cases} \boldsymbol{r}_k^{(t)(1)} \leftrightarrow \boldsymbol{r}_k^{(t+i-1)(1)} \\ \qquad\quad \text{or} \\ \boldsymbol{r}_k^{(t)(2)} \leftrightarrow \boldsymbol{r}_k^{(t+i-1)(2)} \end{cases} \tag{20}$$

```
Algorithm 6 PHY-Layer Authentication With BSDA
Input: The channel matrices of unknown nodes $\boldsymbol{X}_{Auc}$.
Output: The labels of unknown nodes $\hat{\boldsymbol{Y}}_{Auc}$.
1: Get the original CSI dataset $\boldsymbol{D}_{train} = \{\boldsymbol{X}_{train}, \boldsymbol{Y}_{train}\}$ and $\eta$.
2: ECN generates new CSI training dataset $\boldsymbol{D}_{BSDA\_train}$ according to equation (19), (20), (21), (22) and (23).
3: ECN uses the new training dataset $\boldsymbol{D}_{BSDA\_train}$ by performing the functional module of offline training as given in Section IV-A.
4: Obtain $\boldsymbol{X}_{Auc}$.
5: $\boldsymbol{X}_{Auc}$ is fed into the authentication model.
6: Output the labels of unknown nodes $\hat{\boldsymbol{Y}}_{Auc}$.
```

where i is just a positive integer index here and $i = 2, \ldots, \eta$.

Then, the new channel response matrix can be expressed as

$$\begin{aligned}\boldsymbol{\Phi}_k^{\xi} &= \left\{\Phi_k^1, \Phi_k^2, \ldots, \Phi_k^{(\psi_k-\eta+1)\times(2\eta-2)}\right\} \\ &= \left\{\begin{bmatrix}\boldsymbol{r}_k^{(i)(1)} \\ \boldsymbol{r}_k^{(t)(2)}\end{bmatrix}_{m\times 2n}, \begin{bmatrix}\boldsymbol{r}_k^{(t)(1)} \\ \boldsymbol{r}_k^{(t+1)(2)}\end{bmatrix}_{m\times 2n}\right\}\end{aligned} \tag{21}$$

where ξ is the index of the newly generated sample and $\xi = 1, 2, \ldots, (\psi_k - \eta + 1) \times (2\eta - 2)$.

After generating the channel response matrix, the new input sample is

$$\begin{aligned}\boldsymbol{X}_{\mathrm{BSDA}\_k} &= \left\{\mathbf{U}_k^t, \boldsymbol{\Phi}_k^{\xi}\right\} \\ &= \left\{U_k^1, U_k^2, \ldots, U_k^{\psi_k}, \Phi_k^1,\right. \\ &\qquad \left.\Phi_k^2, \ldots, \Phi_k^{(\psi_k-\eta+1)\times(2\eta-2)}\right\}.\end{aligned} \tag{22}$$

The new output sample is

$$\begin{aligned}\boldsymbol{Y}_{BSDA\_k} = \left\{I_k^1, I_k^2, \ldots, I_k^{\psi_k},\right. \\ \left. I_k^1, I_k^2, \ldots, I_k^{(\psi_k-\eta+1)\times(2\eta-2)}\right\}.\end{aligned} \tag{23}$$

The new training data set of the kth CD can be denoted as follows:

$$\boldsymbol{D}_{BSDA\_k} = \left\{\boldsymbol{X}_{BSDA\_k}, \boldsymbol{Y}_{BSDA\_k}\right\}. \tag{24}$$

The PHY-layer authentication with BSDA is summarized in Algorithm 6.

VII. Experimental Results

A. Overall Settings

Two sets of experiments are conducted by using the channels generated by computer simulation and an NI USRP testbed, respectively. A number of counterparts are considered for the comparison purpose, including the case without using TL, the case without data augmentation, and the case using conventional threshold-based PHY-layer authentication, respectively. In addition to TP-Net, we also examine a conventional CNN architecture [32], GC-Net [15], and VGG [24]. Furthermore, the threshold-based method in [8] is implemented, too.

The TP-Net is implemented with four convolutions, among which two are composed of two convolution layers with 1 × 1 filter and 16 and 32 feature maps, respectively; and the other two are composed of two convolution layers with 3 × 3 filters and 32 feature maps. The 2 × 2 max-pooling layer with a 2 × 2 stride as applied after both of the two 3 × 3 convolution layers. GAP is applied to the output of the first convolution layer and the collected parameters are fed as input to the SoftMax layer for classification.

The conventional CNN architecture, namely, CNN-2, employs ReLU as the activation function and contains two convolutions, where one is composed of 4 × 4 filter and eight feature maps, and the other is composed of 2 × 2 filters and 16 feature maps. The 4 × 4 average pooling layer with a stride of 4 × 4 is applied after the 4 × 4 convolution layers. The 2 × 2 average pooling layer with a stride of 2 × 2 is applied after the 2 × 2 convolution layers.

The GC-Net considered in the experiment employs ELU as the activation function and is composed of three convolution layers with small 3 × 3 filters and 64, 64, and 64 feature maps, respectively. The 2 × 2 max-pooling layer with a stride of 2 × 2 is applied after both of the first two convolution layers. GAP is applied to the output of each convolution layer and the collected averaged features are fed as input to a dense layer with 64 neurons. The output of the dense layer is fed as input to the SoftMax layer for classification.

The VGG architecture considered in the experiment, called VGG-7, employs ELU as the activation function and is composed of seven convolution layers with small 3 × 3 filters and 64, 64, 128, 128, 256, 256, and 256 feature maps, respectively. The 2 × 2 max-pooling layer with a stride of 2 × 2 is applied after the first two, the first four and the last convolution layer, respectively. With two fully connected layers, one has 512 neurons followed by an ELU activation function, while the other has "classes" neurons accompanied with the SoftMax activation function, where classes is the total number of neurons to be classified. The adaptive moment estimation (Adam) [33] accelerated gradient algorithm is used for the acceleration of all CNN training. To be fair, the parameter α of ELU is set to 1.

The *authentication rate* is considered as the performance metric, denoted as *AucRate*, which is the ratio of correctly authenticated samples to the total launched ones

$$\mathrm{AucRate} = \frac{1}{\varPsi}\sum_{\varPsi}\sum_{\varOmega}\left(\hat{\boldsymbol{Y}}_{\mathrm{Auc}} \circ \boldsymbol{Y}_{\mathrm{Auc}}\right) \tag{25}$$

where $\varPsi$ is the total number of CSI that need to be verified, $\varOmega$ is the number of nodes, and $\hat{\boldsymbol{Y}}_{\mathrm{Auc}} \circ \boldsymbol{Y}_{\mathrm{Auc}}$ denotes the Hadamard product of the matrices $\hat{\boldsymbol{Y}}_{\mathrm{Auc}}$ and $\boldsymbol{Y}_{\mathrm{Auc}}$. $\hat{\boldsymbol{Y}}_{\mathrm{Auc}}$ is the label of output of CNN, while the real label of the output is $\boldsymbol{Y}_{\mathrm{Auc}}$.

Training time is used to measure the computational complexity of the model training, which is described in (26)

$$\text{Training time} = \sum_{i=1}^{N} t_i \tag{26}$$

where t_i denotes the training time of the ith epoch, i is the number of epoch here, and $i = 1, 2, \ldots, N$.

In addition, the true-positives rate (TPR) and false-positive rate (FPR) are two performance matrices defined as follows:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{27}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TP}} \tag{28}$$

where true positive (TP) and false positive (FP) are the number of malicious CSIs being detected and legitimate CSIs detected as malicious, respectively, and false negative (FN) is the number of malicious CSIs launched without being detected. TP + FN depicts the total number of illegal CSIs, while FP + TP is the total number of CSIs determined as illegal.

We also define the *detection rate*, denoted as *DetRate* and given by

$$\mathrm{DetRate} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{29}$$

where true negative (TN) represents the number of legitimate CSIs being correctly detected.

B. Training With Simulated Channels

The tapped delay line (TDL) model is exploited to simulate the Rayleigh fading channel [34] with multipath delay, where a set of unit-power, independent nonfrequency selective fading generators, such as the filtered white Gaussian noise (FWGN) or Jakes model [35]–[37], are employed. It is based on the following equation:

$$y(n) = \sum_{d=0}^{N_D-1} h_d(n)x(n-d) \tag{30}$$

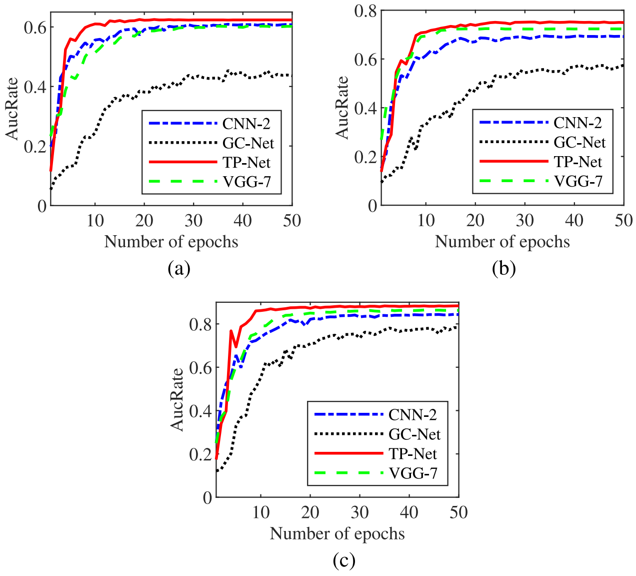
where $y(n)$ and $x(n)$ are the output and input at the nth sample instant, respectively, and N_D denotes the number of taps of the channel filters with filter coefficients h_d. We use five paths with different power delays to synthesize the channels of each node. The time delay of the first four paths is the same, which is 0 s (s), 2×10^{-6} s, 4×10^{-6} s, 8×10^{-6} s, respectively. When there are 40 CDs, the time delay of the fifth path is 3×10^{-6} s, 4×10^{-6} s, $\ldots, 4.2 \times 10^{-5}$ s, respectively.

The least squares (LSs) algorithm is adopted to estimate CSI under OFDM with a sampling interval $t_{\text{sampling}} = 1 \times 10^{-6}$ s, the number of subcarriers $n_{\text{subcarrier}} = 128$, the pilot interval $n_{\text{pilot\_interval}} = 3$, the cyclic prefix length $l_{\text{cp\_length}} = 16$, and the digital modulation method as QPSK. The number of original channel samples for each node training and that for testing CNN are both 100.

Our simulation program runs on a 64-bit Win10 Professional system with an Intel Core i7-9750H of the main frequency 2.59 GHz and the physical memory 16 GB. The Python Keras library is used to build the network.

1) Homogeneous Network Scenario: With the homogeneous network scenario, the number of nodes for the online training and offline training is identical and, thus, there is no need to perform online fine tuning during the processes of offline training and online migration authentication.

Fig. 3 shows the authentication rates under 20 CDs with different channel SNR and the number of antennas, and Fig. 4 demonstrates the results of authentication rate under different numbers of nodes with SNR as 2 dB and eight antennas, where the proposed scheme is compared to the case of using CNN-2, GC-Net, and VGG-7. We first find that increasing the number of antennas and channel SNR can improve the authentication rate, while the latter is even more dominant. In addition, the proposed TL-PHA scheme can achieve the best performance, and such an advantage remains when the number of nodes increases.



(a)

(b)

(c)

Fig. 3. Authentication rate under different conditions. (a) 20 nodes, 2 dB, eight antennas. (b) 20 nodes, 2 dB, 16 antennas. (c) 20 nodes, 8 dB, eight antennas.

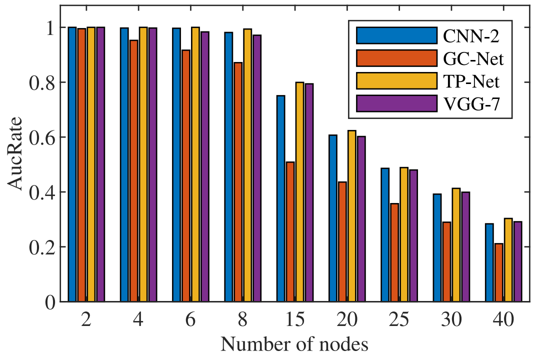


Fig. 4. Authentication rate of each scheme with channel SNR as 2 dB and eight antennas, where the authentication models of all schemes are the training results of the 50th epoch.

Fig. 5 shows the comparison results among all the ML-based and the threshold-based schemes, where the former provides much better performance than the latter, while the performance of threshold-based schemes highly relies on the channel SNR and threshold values. It confirms that the proposed TL-PHA scheme takes all the advantages against its counterparts.

Table I shows the detection performance with channel SNR as 2 dB and number of antennas as 8, where 50% of the nodes in the network are malicious. It is clear the proposed scheme outperforms all the other considered counterparts.

2) Disparate Network Scenario: With the disparate network scenario, the number of online authentication nodes

TABLE I
DETECTION PERFORMANCE UNDER DIFFERENT NODES WITH CHANNEL SNR AS 2 dB AND NUMBER OF ANTENNAS AS 8

Scheme / Nodes	DetRate				TPR				FPR			
	CNN-2	GC-Net	TP-Net	VGG-7	CNN-2	GC-Net	TP-Net	VGG-7	CNN-2	GC-Net	TP-Net	VGG-7
4	0.995	0.94	1	0.988	0.99	0.93	1	0.98	0	0.051	0	0.005
10	0.913	0.904	0.942	0.931	0.94	0.916	0.97	0.962	0.108	0.105	0.081	0.094
20	0.743	0.776	0.802	0.741	0.793	0.821	0.856	0.819	0.279	0.247	0.228	0.292
30	0.678	0.675	0.696	0.68	0.76	0.725	0.795	0.742	0.347	0.341404	0.336	0.339
40	0.625	0.635	0.662	0.629	0.683	0.676	0.679	0.681	0.388	0.375	0.343	0.383

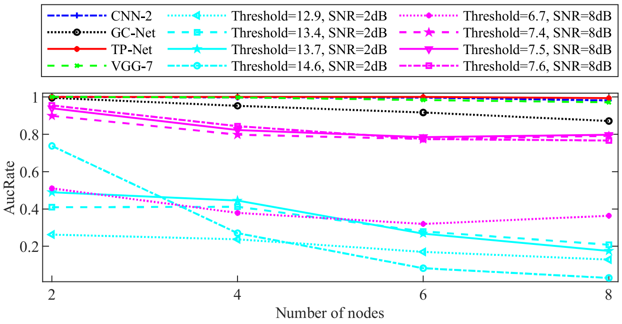


Fig. 5. Comparison of authentication rate under different numbers of nodes among all the ML-based schemes and the threshold-based scheme, where the authentication rate of ML-based schemes is the result of the 50th epoch.

is different from that of the offline training nodes and, thus, the fine tuning of the online authentication model is necessary. Without loss of generality, we focus on the case, where the online authentication nodes are less than that for the offline training. Specifically, the proposed offline training is conducted on a 20-node CSI data set with 100 CSIs of each node. The resultant authentication rates are compared to that without TL under 8-dB channel SNR and a number of eight antennas.

Figs. 6 and 7 demonstrate the authentication rate performance between online TL and training from the scratch among all the ML-based schemes for 15 nodes with 5 and 10 training samples per node, respectively, where "TL, no Freeze" stands for the case where no parameter is frozen, and only the model trained offline is directly transferred for online fine tuning; "TL, Freeze 1C" is the case where the parameters of the first convolution layer of the migrated model are frozen during online fine tuning; "TL, Freeze 2C" is the case where the parameters of the first two convolution layers of the migrated model are frozen, and so on; with "TL, Freeze 1-D," the parameters of the first "Dense" layer and all the convolution layers of the migrated model are frozen during the online fine tuning; and with "no TL," the model is trained from the scratch with 5 or 10 samples per node.

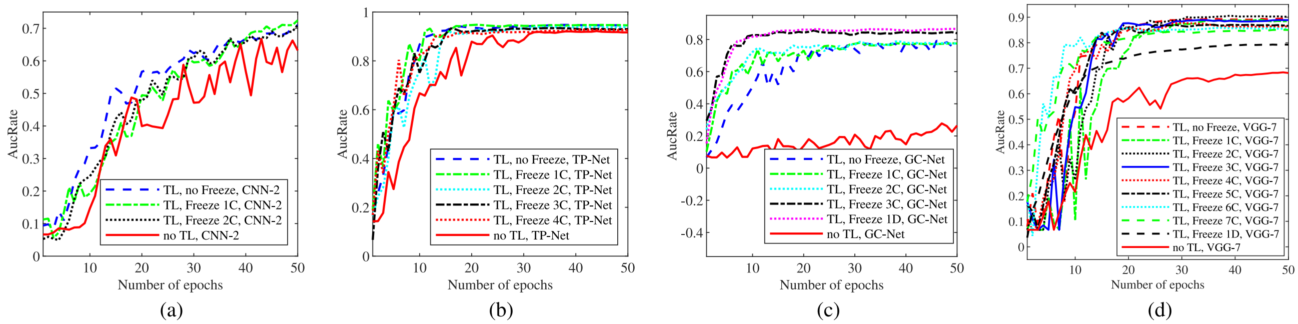
Obviously, the authentication rate performance by using online TL is better than that of the no-TL case. In addition, the authentication rate curve of Fig. 7 is smoother than that of Fig. 6 since the number of training data sets is different. This also shows that with more data sets, the network is better trained. What is more, freezing the parameters of different layers affect the results. For the TL of CNN-2, little difference on authentication rate between the cases of frozen parameters and unfrozen parameters, respectively, is observed. For the TP-Net, the authentication rate of no freezing parameters or freezing the parameters of the first three convolution layers is higher than that of freezing all parameters. For the GC-Net, the authentication rate performance by freezing the parameters of the first three convolution layers and/or the parameters of the first "Dense" layer is better than the other cases, while the authentication rate of using TL is the lowest when the parameters are not frozen. We can see that the use of migration learning is very helpful to improve the authentication rate for GC-Net especially when the amount of training data sets is limited. For the VGG-7, the authentication rate of no freezing parameters or freezing the parameters of the first three convolution layers is higher than that of other cases, while the authentication rate of TL is the lowest when the parameters of dense layer are frozen. This is due to the fact that the parameters of the dense layer directly affect the authentication results of the model, so the parameters of the dense layer should not be frozen when the model is fine tuned.

Fig. 8 shows that the CNN-2 scheme takes the shortest training time, and the proposed TL-PHA scheme consumes slightly higher training time, which demonstrates a graceful compromise with the gained authentication rate performance.

Along with Figs. 6–8, we conclude that using TL can not only improve the authentication rate but also save the training time, which perfectly fits into the EC-related application scenarios.

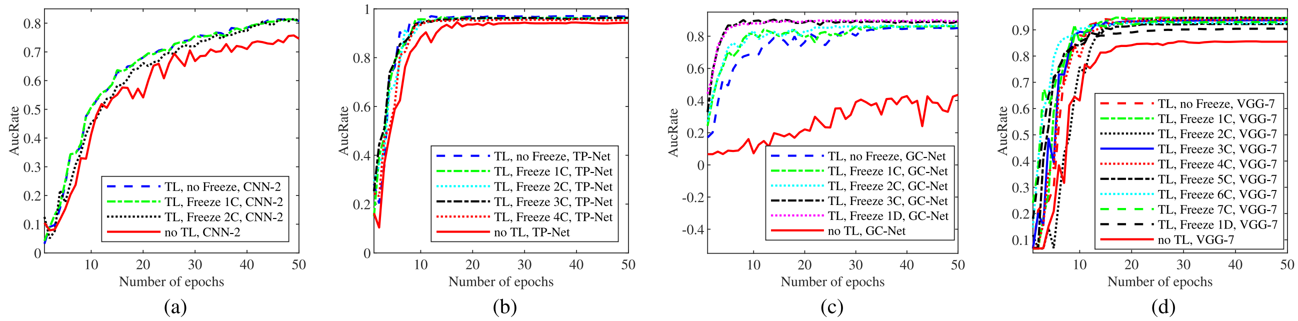
3) Analysis of Data Augmentation: The data augmentation methods are examined on a network by taking 100 original channel data frames in the training of the CNN of each node, where the channel SNR and the number of antennas is 2 dB and 8, respectively. With the SWDA method, the random number is generated by the normal distribution function $\mathcal{N}(\mu, \sigma^2)$ and the uniform distribution function $\mathcal{U}(a, b)$, respectively.

Fig. 9 shows the authentication rate by using the proposed data augmentation methods on all the ML-based schemes for 15 nodes. Clearly, all the schemes with BSDA can yield better authentication rates while with a higher convergence speed. Such an advantage is gained thanks to the fact that BSDA swaps the elements in the same position of the channel matrix instead of the correlation between channel matrices and, thus, can generate more new data samples that facilitate boosting the authentication rate. On the other hand, SWDA using a uniform distribution function can yield a higher authentication rate than the case without data augmentation and the SWDA method using a normal distribution function. This is because the channel noise actually follows a Gaussian distribution and, thus, the corresponding augmented training data sets can be closer to the real channel data. Furthermore, with the increase of θ, the number of newly generated channel samples will



(a) (b) (c) (d)

Fig. 6. Comparison of authentication rate between the TL method and the method of training from the scratch among the ML-based schemes for a 15-node network. The channel SNR is 8 dB and the number of antennas is 8. The number of original channel samples of each node for offline training is 100. The number of new CSI per node of the online fine tuning authentication model is 5. (a) CNN-2. (b) TP-Net. (c) GC-Net. (d) VGG-7.



(a) (b) (c) (d)

Fig. 7. Comparison of authentication rate between the cases of using TL and training from the scratch for 15 nodes. The channel SNR is 8 dB and the number of antennas is 8. The number of original channel samples of each node for offline training is 100. The number of new CSI per node of the online fine tuning authentication model is 10. (a) CNN-2. (b) TP-Net. (c) GC-Net. (d) VGG-7.

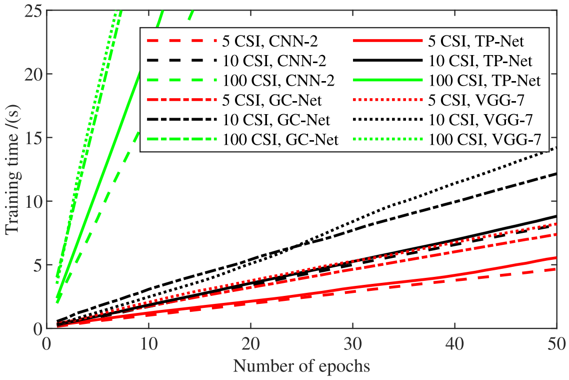


Fig. 8. Comparison of training time of each scheme under 15 wireless nodes and different number of training samples with channel SNR as 8 dB and number of antennas as 8.

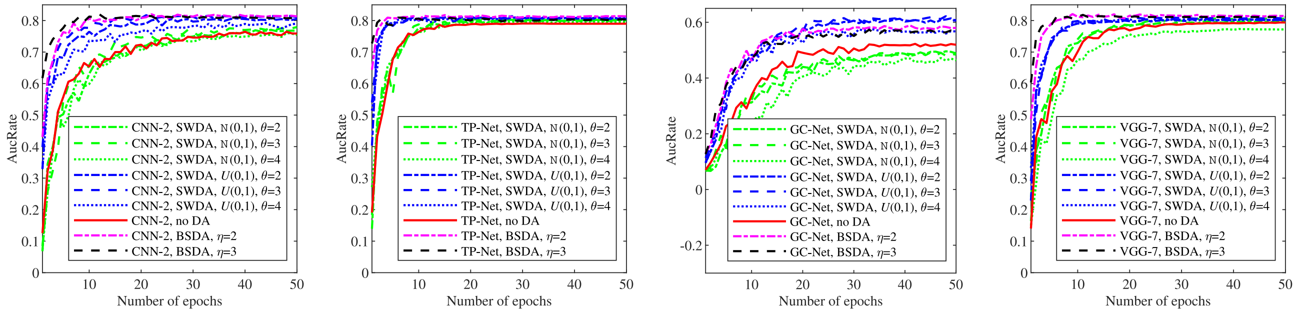
decrease but the authentication rate does not changed much, which is still higher than that without data augmentation.

C. Experiments on NI USRP Platform

Experiments are performed on an NI USRP Platform in an office room of 8-m long, 7.5-m wide, and 3-m high. The ECN is simulated by four USRPs, which configure eight transmit and receive antennas, respectively. Four CDs are simulated accordingly, where one is equipped with two transmit and two receive antennas, while the others are simulated by two USRPs, which configure four transmit and four receive antennas, respectively. The number of original training and testing frames per CD is 100, respectively. Multiple input and multiple output-orthogonal frequency-division multiplexing (MIMO-OFDM) at the ECN is assumed, and the improved-scaled LSs (ILSs) is adopted for channel measurement [28], [29]. The center frequency is $f_c = 3.5$ Giga Hertz (GHz), the number of subcarriers is $n_{\text{subcarrier}} = 128$, the sampling interval $t_{\text{sampling}} = 5 \times 10^{-7}$ s, the digital modulation method is 4QAM, and the transmitting power is 15 dBm and transmission gain 20 dB. The server parameters of training CNN are as follows: the server CPU is with Intel Xeon Silver 4114 with the main frequency as 2.2 GHz, the physical memory as 15 982 940 kB, the operating system as Ubuntu 18.04.4 LTS.

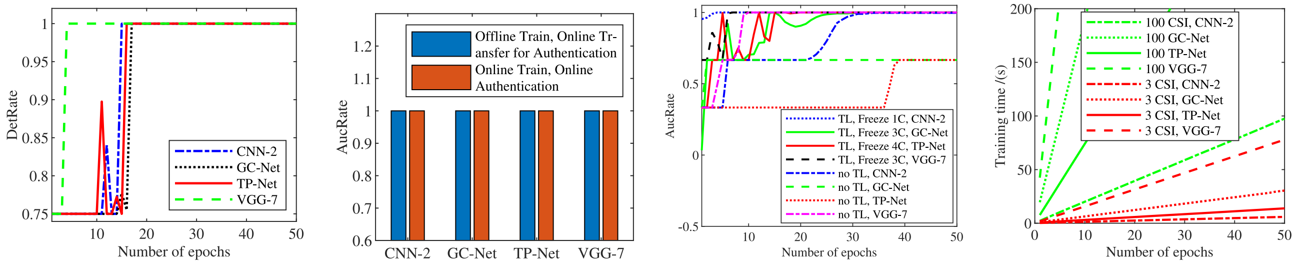
Fig. 10 shows the authentication and detection performance, where the notations follow Figs. 6 and 7. Fig. 10(a) shows that after several epochs, the detection ability of all schemes converges and the detection results are perfect, due to the small number of malicious and legitimate nodes.

From the results of Fig. 10(b), offline training and online training have no effect on the authentication rate. Fig. 10(c) demonstrates the authentication rate between the methods of online TL and training from the scratch among all the ML-based schemes for four nodes with three training samples per node. Obviously, the authentication rates of all schemes are perfect after sufficient training epochs, while by using the online TL scheme, the least amount of training epochs is needed, thanks to much less parameters to be fine tuned online. Fig. 10(d) shows that the consumed training time highly depends on the size of training data sets, and the training time by the TP-Net scheme is slightly longer than that



(a) (b) (c) (d)

Fig. 9. Comparison of authentication rate for different configurations of data augmentation methods. The number of nodes is 15, the channel SNR is 2 dB, and the number of antenna is 8. (a) CNN-2. (b) TP-Net. (c) GC-Net. (d) VGG-7.



(a) (b) (c) (d)

Fig. 10. Authentication and detection performance on the USRP testbed. (a) Comparison of detection rate. There are four nodes in total, including one malicious node and three legitimate nodes. (b) Comparison of authentication rate between online migration authentication of offline training and online authentication of online training for four nodes. (c) Comparison of authentication rate between the method of TL and the method of training from scratch for four nodes. The number of original channel samples of each node for offline training is 100. The number of new CSI per node for the online fine tuning authentication model is 3. (d) Comparison of training time of each scheme under four nodes and different number of training samples.

by CNN-2 but much shorter than that of VGG-7 and GC-Net. This further attests the advantage of using TL in online applications such as the PHY-layer authentication considered in the study.

VIII. Conclusion

In this work, we introduced TL-PHA for lightweight user authentication for latency-sensitive applications such as the EC. It is featured by incorporating a number of unique designs, including the TP-Net, which is a novel CNN architecture jointly employing max, global, and average pooling; the SWDA and BSDA, which are two novel data augmentation algorithms; as well as the TL mechanism, which enables migration of offline training results and fine tuning of network models for the desired online classification purpose. Extensive experiments were conducted by using both computer simulation and an NI USRP testbed. The results showed that the proposed TL-PHA significantly outperforms all the other counterparts in terms of authentication accuracy and the ability of identifying the malicious nodes, while taking the least computational complexity in the model training phase.

References

[1] S. N. Shirazi, A. Gouglidis, A. Farshad, and D. Hutchison, "The Extended cloud: Review and analysis of mobile edge computing and fog from a security and resilience perspective," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2586–2595, Nov. 2017.

[2] R.-F. Liao *et al.*, "Multiuser physical layer authentication in Internet of Things with data augmentation," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2077–2088, Mar. 2020.

[3] L. U. Khan, I. Yaqoob, N. H. Tran, S. M. A. Kazmi, T. N. Dang, and C. S. Hong, "Edge-computing-enabled smart cities: A comprehensive survey," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10200–10232, Oct. 2020.

[4] M. Aazam, S. Zeadally, and K. A. Harras, "Health fog for smart healthcare," *IEEE Consum. Electron. Mag.*, vol. 9, no. 2, pp. 96–102, Mar. 2020.

[5] S.-C. Hung, X. Zhang, A. Festag, K.-C. Chen, and G. Fettweis, "Vehicle-centric network association in heterogeneous vehicle-to-vehicle networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5981–5996, Jun. 2019.

[6] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe, "Using the physical layer for wireless authentication in time-variant channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2571–2579, Jul. 2008.

[7] H. Wen, Y. F. Wang, X. P. Zhu, J. Q. Li, and L. Zhou, "Physical layer assist authentication technique for smart meter system," *IET Commun.*, vol. 7, no. 3, pp. 189–197, Feb. 2013.

[8] F. Pan *et al.*, "Authentication based on channel state information for industrial wireless communications," in *Proc. 44th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2018, pp. 4125–4130.

[9] F. Pan *et al.*, "Threshold-free physical layer authentication based on machine learning for industrial wireless CPS," *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6481–6491, Dec. 2019.

[10] H. Fang, X. Wang, and L. Hanzo, "Learning-aided physical layer authentication as an intelligent process," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2260–2273, Mar. 2019.

[11] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10037–10047, Dec. 2016.

[12] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[13] N. Wang, T. Jiang, S. C. Lv, and L. Xiao, "Physical-layer authentication based on extreme learning machine," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1557–1560, Jul. 2017.

[14] H. Fang, X. Wang, and S. Tomasin, "Machine learning for intelligent authentication in 5G and beyond wireless networks," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 55–61, Oct. 2019.

[15] Z. Chen and P. Ho, "Cloud based content classification with global-connected net (GC-Net)," in *Proc. 21st Conf. Innovat. Clouds Internet Netw. Workshops (ICIN)*, Feb. 2018, pp. 1–6.

[16] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, p. 60, 2019.

[17] Y. Li, H. Hu, and G. Zhou, "Using data augmentation in continuous authentication on smartphones," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 628–640, Feb. 2019.

[18] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," *IEEE Signal Process. Lett.*, vol. 24, no. 3, pp. 279–283, Mar. 2017.

[19] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, Dec. 2018.

[20] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[21] L. Shao, F. Zhu, and X. Li, "Transfer learning for visual categorization: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 5, pp. 1019–1034, May 2015.

[22] R. Nassif, S. Vlaski, C. Richard, J. Chen, and A. H. Sayed, "Multitask learning over graphs: An approach for distributed, streaming machine learning," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 14–25, May 2020.

[23] G. Oligeri, S. Raponi, S. Sciancalepore, and R. Di Pietro, "PAST-AI: physical-layer authentication of satellite transmitters via deep learning," 2020. [Online]. Available: arXiv:2010.05470.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1097–1105, Jan. 2012.

[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014. [Online]. Available: arXiv:1409.1556.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2016, pp. 770–778.

[27] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, 2015, pp. 1–9.

[28] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.

[29] E. G. Larsson, G. Q. Liu, J. Li, and G. B. Giannakis, "Joint symbol timing and channel estimation for OFDM based WLANs," *IEEE Commun. Lett.*, vol. 5, no. 8, pp. 325–327, Aug. 2001.

[30] Y. Liu, Z. Tan, H. Hu, L. J. Cimini, and G. Y. Li, "Channel estimation for OFDM," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1891–1908, 4th Quart., 2014.

[31] A. Sorrentino, G. Ferrara, and M. Migliaccio, "On the coherence time control of a continuous mode stirred reverberating chamber," *IEEE Trans. Antennas Propag.*, vol. 57, no. 10, pp. 3372–3374, Oct. 2009.

[32] R.-F. Liao *et al.*, "Deep-learning-based physical layer authentication for industrial wireless sensor networks," *Sensors*, vol. 19, no. 11, p. 2440, May 2019.

[33] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. [Online]. Available: arXiv:1412.6980.

[34] R. Liao, H. Wen, J. Wu, H. Song, F. Pan, and L. Dong, "The rayleigh fading channel prediction via deep learning," *Wireless Commun. Mobile Comput.*, vol. 2018, Jul. 2018, Art. no. 6497340.

[35] R.-F. Liao *et al.*, "Security enhancement for mobile edge computing through physical layer authentication," *IEEE Access*, vol. 7, pp. 116390–116401, Aug. 2019.

[36] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Techn. J.*, vol. 47, no. 6, pp. 957–1000, Jul.-Aug. 1968.

[37] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*. New York, NY, USA: Wiley-IEEE Press, 1994.

Yi Chen (Student Member, IEEE) received the master's degree in communication and information systems from Guizhou University, Guiyang, China, in 2016. He is currently pursuing the Ph.D. degree in communication and information system with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu, China, under the supervision of Prof. H. Wen.

He was a visiting Ph.D. student with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, from January 2020 to January 2021. His research interests include communication network, Internet of Things, and physical-layer authentication.

Pin-Han Ho (Fellow, IEEE) received the Ph.D. degree from Queen's University, Kingston, ON, Canada, in 2002.

He is a Full Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. He has authored/coauthored of more than 400 referred technical papers, several book chapters, and coauthored of two books on optical Internet design and optimization. His current research interests cover a wide range of topics in broadband wired and wireless communication networks, including survivable network design, wireless communications, cyber–physical systems, and Internet of Things.

Hong Wen (Senior Member, IEEE) received the M.Sc. degree in electrical and computer engineering from Sichuan University, Chengdu, China, in 1997, the first Ph.D. degree in electrical and computer engineering from Southwest Jiaotong University, Chengdu, China, in 2004, and the second Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2018.

She is currently a Professor with the School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu. Her current research interests include communication systems, physical-layer security, smart grid, and industrial communications.

Shih Yu Chang (Senior Member, IEEE) received the B.S.E.E. degree from National Taiwan University, Taipei, Taiwan, in 1998, and the Ph.D. degree in electrical engineering and computer engineering from the University of Michigan, Ann Arbor, MI, USA, in 2006.

From August 2006 to February 2016, he was the Faculty with the Department of Computer Engineering, National Tsing Hua University, Hsinchu, Taiwan. From July 2007 to August 2007, he had been a Visiting Assistant Professor with Television and Networks Transmission Group, Communications Research Centre, Ottawa, ON, Canada. In June 2018, he began to provide lectures about machine learning, data science, and AI with San Jose State University, San Jose, CA, USA. Besides academic position, he also works as an AI technical lead focusing on applying machine learning techniques to automate office works.

Shahriar Real received the M.A.Sc. degree in pattern analysis and machine intelligence from the University of Waterloo, Waterloo, ON, Canada, in 2020.

Alongside completing his degree, he has two years of intensive research experience working as a Graduate Research Student with the University of Waterloo. He is currently working with KPMG, Amstelveen, The Netherlands, as a Senior Software Consultant in their Scientific Research and Experimental Development Team. His research interest lies in machine learning, neural networks, and software methodologies.