# Manchester Metropolitan University

Deng, Dan and Li, Xingwang and Zhao, Ming and Rabie, Khaled and Kharel, Rupak (2020) Deep Learning Based Secure MIMO Communications with Imperfect CSI for Heterogeneous Networks. Sensors, 20 (6). ISSN 1424-8220

https://e-space.mmu.ac.uk

# Deep Learning-Based Secure MIMO Communications with Imperfect CSI for Heterogeneous Networks

**Dan Deng [1] , Xingwang Li [2,*], Ming Zhao [3], Khaled M. Rabie [4] and Rupak Kharel [5]**

[1] School of Information Engineering, Guangzhou Panyu Polytechnic, Guangzhou 511406, China; dengdan@ustc.edu

[2] School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China

[3] CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China, Hefei 230027, China; zhaoming@ustc.edu.cn

[4] Department of Engineering, Manchester Metropolitan University, Manchester M15 6BH, UK; k.rabie@mmu.ac.uk

[5] Department of Computing and Mathematics, Manchester Metropolitan University, Manchester M15 6BH, UK; r.kharel@mmu.ac.uk

**\*** Correspondence: lixingwang@hpu.edu.cn;

**Abstract:** Perfect channel state information (CSI) is required in most of the classical physical-layer security techniques, while it is difficult to obtain the ideal CSI due to the time-varying wireless fading channel. Although imperfect CSI has a great impact on the security of MIMO communications, deep learning is becoming a promising solution to handle the negative effect of imperfect CSI. In this work, we propose two types of deep learning-based secure MIMO detectors for heterogeneous networks, where the macro base station (BS) chooses the null-space eigenvectors to prevent information leakage to the femto BS. Thus, the bit error rate of the associated user is adopted as the metric to evaluate the system performance. With the help of deep convolutional neural networks (CNNs), the macro BS obtains the refined version from the imperfect CSI. Simulation results are provided to validate the proposed algorithms. The impacts of system parameters, such as the correlation factor of imperfect CSI, the normalized doppler frequency, the number of antennas is investigated in different setup scenarios. The results show that considerable performance gains can be obtained from the deep learning-based detectors compared with the classical maximum likelihood algorithm.

**Keywords:** physical-layer security; deep learning; imperfect CSI; heterogeneous networks; channel estimation

## 1. Introduction

In recent years, the heterogeneous wireless network, which can support high-density and high-rate traffic, has attracted much research interest from both academic and industry sectors [1–3]. The cooperation between the macro base stations (BSs) and the femto BSs can greatly improve the quality of service (QoS) of the user equipment, as well as the spectrum efficiency and the energy efficiency [4]. Therefore, ultra-dense heterogeneous technology has been adopted as one of the potential solutions to the next-generation wireless sensor network [5].

On the other hand, the architecture of heterogeneous wireless networks brings up many challenges to the physical-layer security [6–8]. Several signal-level algorithms, such as beamforming [9,10], artificial noise [11,12] or stochastic geometry approach [13,14], have been proposed as the classical

solutions to the secure multiple-input multiple-output (MIMO) communication scheme. Consequently, as an improvement, deep learning-based secure communication has been proposed to protect the message transmission to the legitimate users [15]. However, due to the noise and the signal distortion or delay of channel estimation, it is difficult to obtain the ideal channel state information (CSI) from both the legitimate users and the eavesdroppers in most of practical scenes [16–21]. For example, authors in [22–24] derived a new outage probability expression and optimize the transmission capacity under different assumptions.

Deep learning technology has already shown astonishing capabilities in dealing with complex network in mobile communications, such as channel estimation, signal detection, modulation recognition and channel equalization. For example, Ref. [25] used a 2D image to present the time-frequency channel fading matrix. By using a super-resolution network and a denoising network, the proposed scheme produces more accurate channel estimation. Under the constraint of one-bit quantization, authors in [26] proposed deep neural network-based auto-encoder for OFDM receiver. Modulation recognition algorithm was considered in [27], where a convolutional neural network followed by a long short-term memory as the classifier was adopted to improve the robustness for modulation recognition. A CNN-based channel prediction scheme was designed in [28] in massive MIMO systems under channel aging effects, where an autoregressive network was used to model the temporal channel correlation of the wireless channel. Considering perfect free space optical communications, a pilot independent deep learning-based channel estimator was proposed in [29,30]. Simulation results indicated that the proposed scheme provided close enough performance to the perfect channel estimation scheme.

Although there are some works on learning-based channel estimation algorithms, with imperfect CSI, the impacts of CNN-based MIMO detector on physical-layer security performance is still an open question. Motivated by that, in this paper, we will discuss the deep learning-based secure MIMO communication algorithm for heterogeneous networks with imperfect CSI. In the heterogeneous networks, there exist both macro BSs and femto BSs, which can serve a group of users within different areas. In general, the macro BS provides the massive connections and the large-scale cell coverage in the hot spots. The femto BS can help enlarge the wireless coverage and the data rates of the user in network edge. In some practical scenarios, secure messages transmitted from the macro BS to the users may be intercepted by the femto BS. In this case, with the help of zero-forcing algorithm [31], the macro BS chooses null-space eigenvectors to prevent information leakage to the femto BS. Classically, the receiver obtains the CSI through pilots signal transmitted from the base station, and there exists a time difference between the channel estimation and the data packet transmission. Thus, the estimated CSI is an imperfect version of the instant of packet transmission, and the zero-forcing secure MIMO algorithm should be re-designed. In this paper, two deep CNN-based detectors are proposed, and a training set with imperfect CSI as well as the original messages or ideal CSI is fed to the CNN model to produce the refined CSI. Simulation results show that the proposed deep learning-based detectors outperform the classical maximum likelihood detector (MLD), especially in small correlation factor cases. The impacts of system parameters, such as the correlation factor, the antenna number, are evaluated in different setup scenarios.

The main contributions of this paper are as follows:

- We employ the deep learning-based technique for secure MIMO communications in heterogeneous networks, which can exploit the benefits of CNN learning model to produce more accurate CSI and meanwhile reduce the bit error rate (BER) of the receiver.
- We provide the detailed framework of deep learning-based detectors, where imperfect CSI as well as the original messages or ideal CSI are included in the training set, and can be used in different application scenarios.

- We present simulation results for deep learning-based detectors in heterogeneous networks. With the help of the CNN technique, the proposed detectors show obvious performance gain over the MLD with acceptable computational cost.

*Notations:* We use $\mathcal{CN}(\mu, \sigma^2)$ to represent the circularly symmetric complex Gaussian random variable with mean $\mu$ and variance $\sigma^2$, $f_X(x)$ and $F_X(x)$ denote the probability density function (PDF) and cumulative distribution function (CDF) of a random variance $x$, respectively, $\text{diag}(\mathbf{A})$ is a row vector consisting of all diagonal elements of $\mathbf{A}$, $\mathbf{A}^*$ is the conjugate transpose of the $\mathbf{A}$, and $\mathbf{H}_{FM}$ denotes the wireless channel fading matrix from $M$ to $F$.

## 2. Related Work

Considering Gaussian wiretap channel, Fritschek et al. [32] introduced an auto-encoder to model the noised wireless channel with a novel security loss function. The generative model of the auto-encoder was trained to encode a message such that the eavesdropper cannot decode it correctly. Results show that the proposed scheme learns a trade-off between legitimate communication rate and secrecy capacity. With the help of the channel's statistical characteristic in relay networks, the authors in [33] proposed a new deep learning-based algorithm to design secure beamforming vector. Considering visible light communication, Xiao et al. [34] proposed deep reinforcement learning (DRL)-based secure communications strategy. Since the optimization of the system secrecy rate is non-convex and NP-hard, a suboptimal solution on beamforming vector can be obtained by introducing zero-forcing beamforming gain to the eavesdropper.

With perfect CSI, the deep feedforward neural network (DFNN) with three layers was adopted for time-slot wireless powered system in [35]. In the proposed scheme, the tuple system parameters, such as the time allocation factor, the power allocation factor as well as the rate of the wiretap channel were produced by the DFNN. During the training phase, the output of the DFNN was compared with the optimal system parameters obtained from exhaust search, and the mean squared error (MSE) was adopted as the performance loss function. Numerical results of the DFNN and the optimal parameters were provided to validate the proposed scheme. Using Stackelberg equilibria, the authors in [36] proposed a secure mobile crowd-sensing (MCS) scheme, where the DRL technique was adopted to derive the optimal MCS policy.

## 3. System Model

Figure 1 depicts the model of secure MIMO communications for heterogeneous networks. In the considered system, there is a macro BS, a femto BS and a terminal user. It is assumed that three types of nodes are all equipped with multiple antennas, and the numbers of antennas are denoted as $N_M$, $N_F$ and $N_U$, respectively. In the heterogeneous networks, macro BS and femto BS work cooperatively to provide the wireless coverage. Specifically, the users located in the hot spots area are associated with the macro BS, and the users located in the network edge are served by the femto BS. We use $\mathbf{H}_{FM}$ and $\mathbf{H}_{UM}$ to denote the wireless channel fading matrix from the macro BS to the femto BS and the users, respectively.

In addition, the channel fading matrix is modeled as time-varying and flat fading using the classical Jakes model [37]. Thus, the correlation coefficient between adjacent samples is given as

$$\rho = J_0(2\pi f_d), \tag{1}$$

where $f_d$ is the normalized doppler frequency spread and $J_0(\cdot)$ is the zero-order Bessel function of the first kind.

Thus, the channel fading matrix can be calculated as

$$\mathbf{H}_{UM}(n) = \rho \mathbf{H}_{UM}(n-1) + \sqrt{1-\rho^2} \mathbf{N}_{UM}(n),$$ (2)

where $n$ denotes the sample time and $N_{UM}(n)$ denotes the additional white noise matrix with the same size of $H_{UM}(n)$. The same equation can be applied on $\mathbf{H}_{FM}$ as follows:

$$\mathbf{H}_{FM}(n) = \rho \mathbf{H}_{FM}(n-1) + \sqrt{1-\rho^2} \mathbf{N}_{FM}(n).$$ (3)

Please note that in the following sections, the sample time $n$ is omitted without loss of generality.

It is well-known that the deep learning networks can effectively capture the correlation features of the training data set. Spatial correlation or antenna correlation, which may achieve dimensionality reduction, is an important challenge for MIMO systems [38,39]. Also, it will be a direction of our future work.

Since the time correlated MIMO channel model is adopted in this paper, we can use DCNN to obtain more accurate CSI from the outdated CSI. Actually, considering the phase rotation introduced by the channel matrix, the outdated CSI is necessary to assist the data recovery at the user.
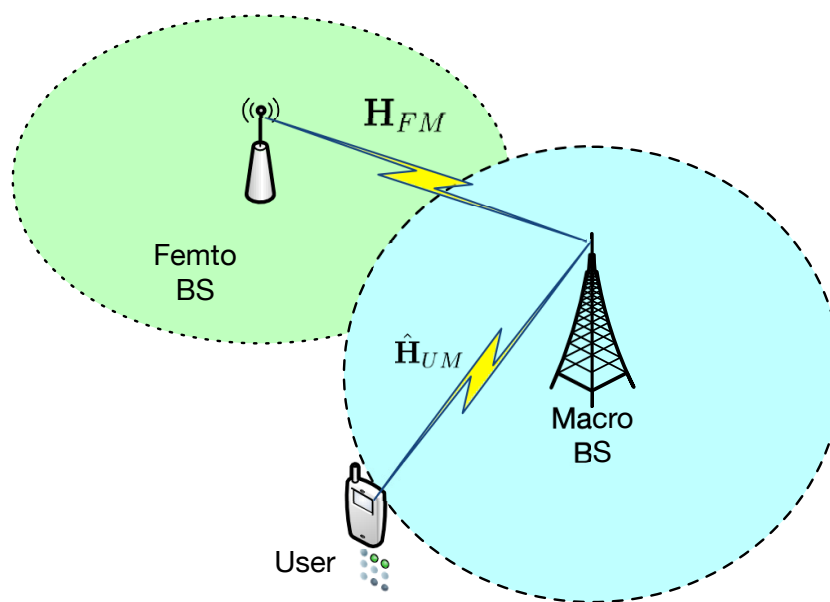


**Figure 1.** Secure MIMO communications for heterogeneous networks.

Due to the open nature of the heterogeneous networks, the femto BS may intercept the signal transmitted from the macro BS to the users. To prevent the information leakage to the femto BS, the macro BS can zero-forcing the equivalent channel matrix of the femto BS by the null-space technique. In this case, the macro BS first obtains the CSI $\mathbf{H}_{FM}$ between the macro BS and the femto BS. Then the null-space eigenvectors can be produced by applying eigenvalue decomposition on the autocorrelation matrix, i.e., $\mathbf{H}_{FM}^*\mathbf{H}_{FM}$, that is

$$(\mathbf{v}, \mathbf{V}) = Eig(\mathbf{H}_{FM}^*\mathbf{H}_{FM}),$$ (4)

where $Eig(\cdot)$ denotes the eigenvalue decomposition, **v** denotes the eigenvalues in ascending order and **V** are the corresponding eigenvectors. Considering the size limitation for both the femto BS and users, it is reasonable to assume that the antenna number of the macro BS is larger than the femto BS and users, i.e., $N_M > N_F, N_M > N_U$. Thus, the number of zero eigenvalues can be given as

$$N_D = N_M - N_F. \tag{5}$$

Please note that $N_D$ is also the number of null-space vectors for $\mathbf{H}_{FM}$, which is given as

$$\mathbf{B} = \mathbf{V}_{:N_D} \tag{6}$$

where $\mathbf{V}_{:N_D}$ denotes the first $N_D$ column vectors of **V**. Thus, the beamforming matrix **B**, which is used by the macro BS to transmit messages to its associated user, lies in the null-space of $\mathbf{H}_{FM}$. Then the signal received at terminal user can be expressed as

$$\mathbf{y} = \sqrt{P}\mathbf{H}_{UM}\mathbf{B}\mathbf{x} + N, \tag{7}$$

where $P$ is the transmission power of the macro BS, $\mathbf{x} \in \mathcal{CN}(0, \mathbf{I})$ is the original message transmitted from macro BS with size $N_D$ and $N \in \mathcal{CN}(0, \sigma^2\mathbf{I})$.

Since the equivalent channel fading matrix of the femto BS is zero-forced, we only need to observe the signal-noise-ratio (SNR) of the associated user. We define the average normalized SNR received at the user as

$$\gamma = \frac{P}{\sigma^2 N_M}. \tag{8}$$

Classically, the receiver obtains the wireless CSI through pilot signals transmitted from the BS, and there exists a time difference between the channel estimation and the data packet transmission. Thus, the estimated CSI is an imperfect version of the instant of packet transmission. To reduce the analysis complexity, it is assumed that the estimation of $\mathbf{H}_{FM}$, while that of $\mathbf{H}_{UM}$ is imperfect. Specifically, the imperfect equation is modeled as

$$\hat{\mathbf{H}}_{UM} = \sqrt{\xi}\mathbf{H}_{UM} + \sqrt{1-\xi}\mathbf{N}_{UM}, \tag{9}$$

where $\xi$ is the correlation factor of the imperfect version of channel matrix.

The standard maximum likelihood detector (MLD) with imperfect CSI can be employed to detect the original message as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\Omega} \|\mathbf{y} - \hat{\mathbf{H}}_{UM}\mathbf{x}\|^2, \tag{10}$$

where $\Omega$ is the all possible constellations set.

## 4. Deep CNN-Based Detector

The imperfect CSI, which is introduced by the noise or delay of channel estimation, will greatly deteriorate the system performance of standard MLD. To overcome the effects of the imperfect CSI in secure MIMO communications, two types of deep CNN (DCNN)-based detectors are proposed in this section, which can be used in different application scenarios. The deep CNN models are first trained with

predefined loss functions and then used to generate the refined CSI $\widetilde{\mathbf{H}}_{UM}$, which can be fed to the MLD to obtain the original message, i.e.

$$\hat{\mathbf{x}}_{cnn} = \arg \min_{\mathbf{x} \in \Omega} \|\mathbf{y} - \widetilde{\mathbf{H}}_{UM}\mathbf{x}\|^2. \tag{11}$$

The details of the DCNN model is given as Figure 2, where there exist $N$ one-dimension convolutional layers excluding the input layer. In the input layer, the channel fading matrix $\hat{\mathbf{H}}_{UM}$ is reshaped as a column vector. Since only real data can be processed in the CNN model, the complex data of $\hat{\mathbf{H}}_{UM}$ can be treated as two real channels [40]. Please note that each convolutional layer is followed by a ReLU activation function except the output layer. Moreover, in the $n$-th convolutional layer, there are $\{F_n, n \in [1, N]\}$ features maps with filter length $\{L_n, n \in [1, N]\}$. Specifically, in the output layer, there is only one feature map.

As to the detailed architecture of the DCNN, we must find the trade-off between complexity and performance. It is noted that the fully connected DNN, which may hold better performance, while its computational complexity is proportional to the square of the number of nodes. On the other hand, both the training data set and the training time required by fully connected DNN the is too large to be satisfied. There also exist some powerful CNN models, such as VGG [41] and ResNet [42], which improve the detection probability by increasing the depth of the models to 19 and 34, respectively. Specifically, the number of parameters for VGG-19 is up to 144M. Thus, to decrease the computational complexity, we must simplify the classical CNN models as follows. The architecture of DCNN model includes $N{=}4$ layers, and is described by $F_n$ and $L_n$ as $F_n = \{32, 16, 8, 1\}$, $L_n = \{36, 3, 3, 36\}$.

In the following sections, two different methods, denoted as DCNN type-I and type-II, are introduced to train the DCNN model.
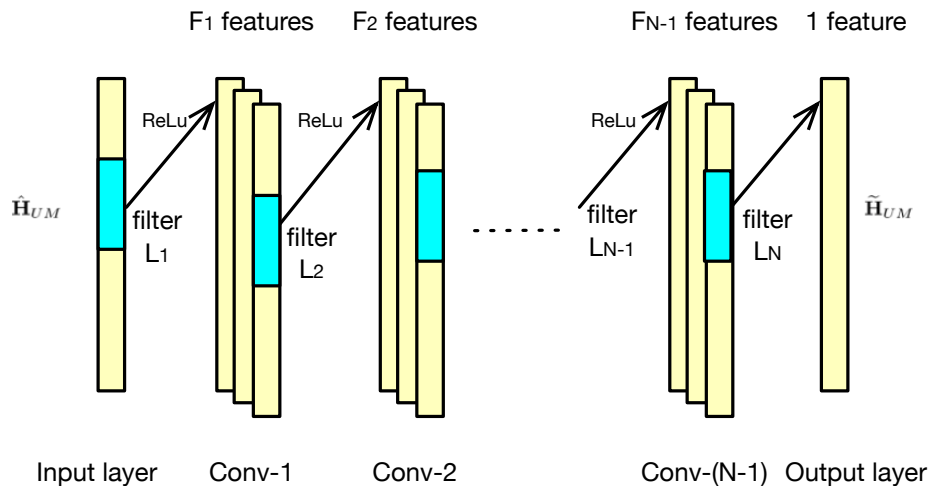


**Figure 2.** Details of DCNN model.

### 4.1. DCNN Type-I: Training with Accurate CSI

Figure 3 shows the first training method of DCNN-based detector with accurate CSI, which is denoted as DCNN type-I. In the training phase, a data set including both the imperfect CSI $\hat{\mathbf{H}}_{UM}$ and the accurate

CSI $\mathbf{H}_{UM}$ is fed into the learning model. The loss function is defined as the mean square error between the output of the model $\widetilde{\mathbf{H}}_{UM}$ and the accurate CSI $\mathbf{H}_{UM}$, i.e.

$$\epsilon_I = \frac{\|\widetilde{\mathbf{H}}_{UM} - \mathbf{H}_{UM}\|^2}{N_U N_M}. \tag{12}$$

During the DCNN model training, the loss function is calculated batch by batch and used to optimize the weight and the bias of the DCNN model [43]. Please note that the accurate CSI is necessary for DCNN type-I, which is used to calculate the model loss function. Thus, the application of DCNN type-I is limited, because in some practical scenarios, it is difficult to obtain the accurate CSI especially in the wireless MIMO communications. Therefore, another type of DCNN is proposed to overcome this limitation.
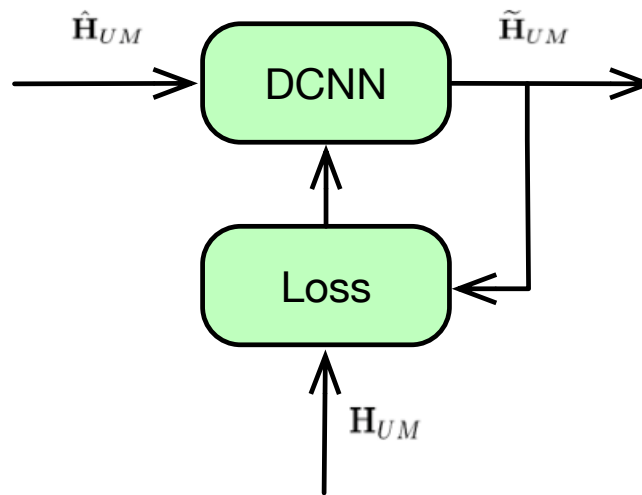


**Figure 3.** DCNN type-I: Training model of CNN with accurate CSI.

*4.2. DCNN Type-II: Training with Original Message*

The training architecture of DCNN type-II is given as in Figure 4, where accurate CSI is not needed. Instead, the output of DCNN $\widetilde{\mathbf{H}}_{UM}$ is used in MLD and obtain the likelihood of each candidate message as follows:

$$\hat{q}_i = \exp\{-\|y - \widetilde{\mathbf{H}}_{UM}\mathbf{B}\mathbf{x}_i\|^2\}, x_i \in \Omega. \tag{13}$$

By using of the SoftMax function, the normalized likelihood probability of each candidate message can be given as

$$q_i = \frac{\hat{q}_i}{\sum_{i=1}^{|\Omega|} \hat{q}_i}. \tag{14}$$

We use $p_i, i \in [1, |\Omega|]$ to denote the correct probability of each candidate message. That is $p_i = 1$ if $i$-th candidate message is correct, otherwise $p_i = 0$. Inspired by the information theory, the cross-entropy can be used to quantify the difference between two probability vectors.

Accordingly, for probability distributions of $q_i$ and $p_i$, we can calculate the cross-entropy as follows:

$$C(p,q) = \sum_{i=1}^{|\Omega|} p_i log_2 \frac{1}{q_i}. \tag{15}$$

Then, the cross-entropy $C(p,q)$ can be used as the loss function to train the DCNN model.

Compared with the loss function in (12), only the received signal $y$, the beamforming matrix **B** and the original message $x$ are needed, which enlarge the application scenarios of the DCNN type-II. On the other hand, without the help of the accurate CSI, DCNN type-II leads to deteriorated performance compared with type-I, which can be validated in the simulation results.

Please note that the DCNN model could also output the ground-true symbol $x$ directly in a supervised manner, and the outdated MIMO channel matrix could be further employed as side information by inputting it to the DCNN. We use DCNN type-III to denote the new DCNN model. Although the detailed architectures of DCNN type-II and the suggested DCNN type-III are different, they are functionally equivalent as a black box with DCNN kernels. In other words, the MLD module in DCNN type-II can be seen as part of the functions of DCNN type-III.
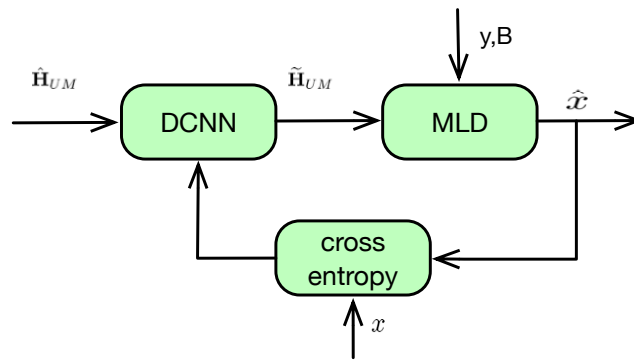


**Figure 4.** DCNN type-II: Training model of CNN without accurate CSI.

## 5. Simulation Results

In this section, simulation results are provided to verify the proposed DCNN models. The impacts of system parameters, such as the correlation factor of imperfect CSI $\xi$, the normalized doppler frequency $f_d$, the number of antennas is evaluated in different setup scenarios. Since the equivalent channel fading matrix of femto BS is zero-forced, BER of users with different detectors is used to evaluate the system performance.

Specifically, QPSK modulation is employed in all simulation setups. Since the constellations of QPSK modulation is 2-D complex signal, we can generalize the setup to an arbitrary modulation order. We set the data packet length as 600 bits, and each batch consists 10 data packets. During the training phase, a training data set with 10000 batches as well as a validation data set with 1000 batches are fed to the DCNN, and a test data set with 1000 batches is used to evaluate the BER performance of the proposed schemes. The popular TensorFlow framework [44] is adopted in our simulations, while the adaptive moment estimation (Adam) optimizer is used to minimize the loss value during training phase. In particular, the optimization parameters are listed as follows: learning rate is 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$. Since the testbed of our paper is under construction at this moment, we will present experiment results on real datasets in future works.

Figure 5 depicts the BER performance versus the SNR of the system with MIMO setup as $N_M = 4, N_U = 4, N_F = 2$, the normalized doppler frequency $f_d = 0.1$, and the correlation factor of imperfect CSI $\xi = 0.90$. The BER of three types of detectors, such as standard MLD, DCNN type-I and DCNN type-II are compared. As a benchmark, the BER curves obtained by MLD with the perfect CSI is also presented. We can see from this figure that the outdated CSI has obvious adverse effects on BER performance. As shown in this figure, in the high SNR region, DCNN-based detectors show a performance gain of about 4dB in comparison to the standard MLD. The reason is that the former can refine the imperfect channel matrix and produce more accurate CSI, then the BER of the system can be improved.

Moreover, two types of DCNN-based detectors show almost the same performance with slight gap. Similar results can be obtained from Figures 6 and 7, where the correlation factors of imperfect CSI are $\xi = 0.8$ and $\xi = 0.7$, respectively. However, accurate CSI is necessary for DCNN type-I, which is used to calculate the model loss function. Thus, the application of DCNN type-I is limited, because in some practical scenarios, it is difficult to obtain the accurate CSI especially in the wireless MIMO communications. Compared with the loss function, only the received signal $y$, the beamforming matrix **B** and the original message $x$ are needed, which enlarge the application scenarios of the DCNN type-II. On the other hand, without the help of the accurate CSI, DCNN type-II leads to deteriorated performance compared with type-I.
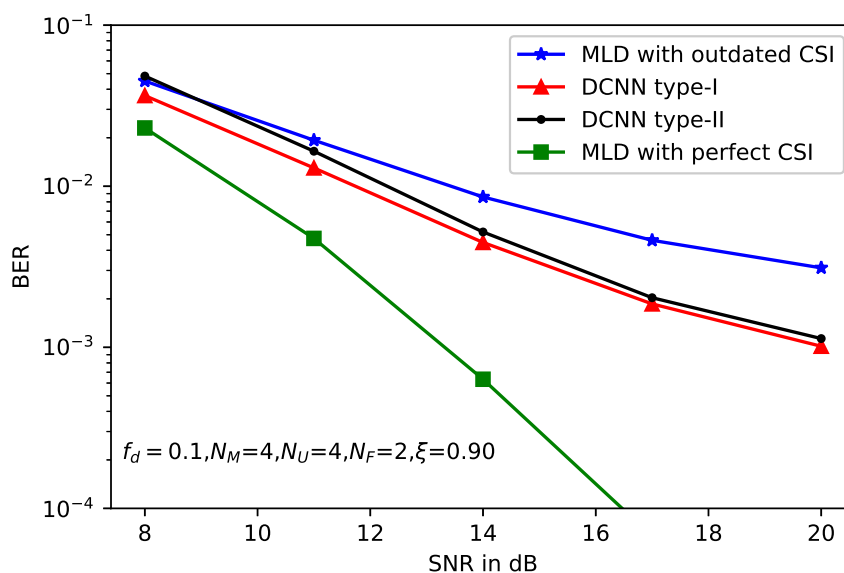


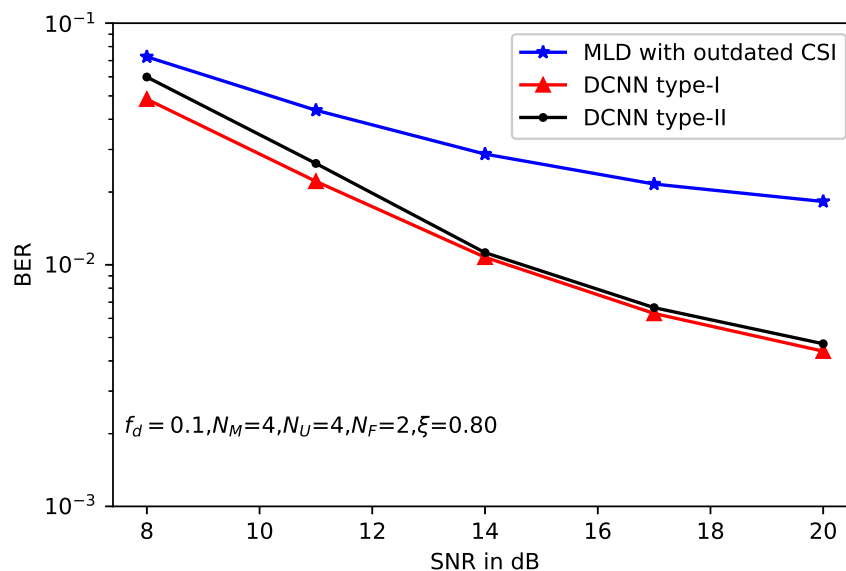**Figure 5.** BER performance of deep learning-based MLD with $\xi = 0.9$.

**Figure 6.** BER performance of deep learning-based MLD with $\xi = 0.8$.
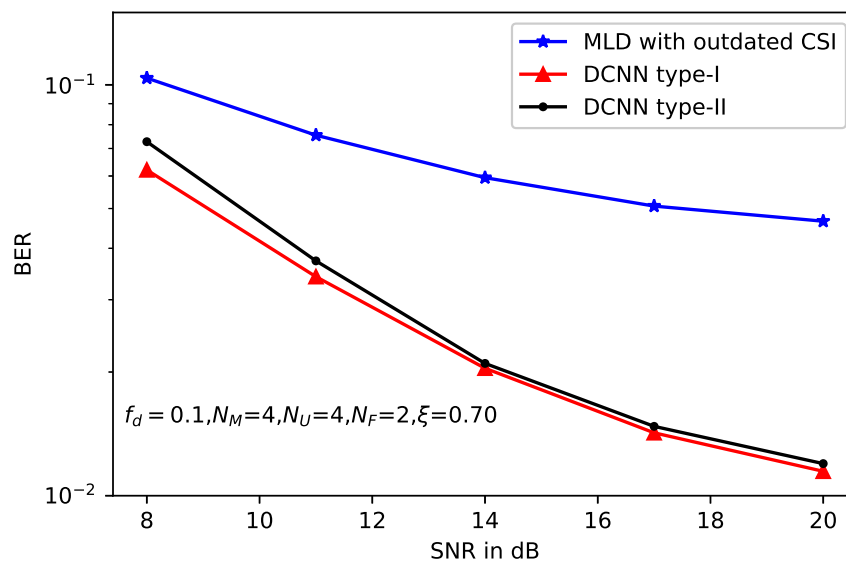


**Figure 7.** BER performance of deep learning-based MLD with $\xi = 0.7$.

Figure 8 depicts the BER performance versus the correlation factor of imperfect CSI $\xi$, where the average SNR of the system is SNR = 20 dB. The BER of three types of detectors, such as standard MLD, DCNN type-I and DCNN type-II are compared. As shown in this figure, DCNN-based detectors indicate considerable performance gains relative to the standard MLD. This is because DCNN-based detectors can refine the imperfect channel matrix and produce more accurate CSI, hence the BER of the system can be improved. Moreover, two types of DCNN-based detectors show almost the same performance.
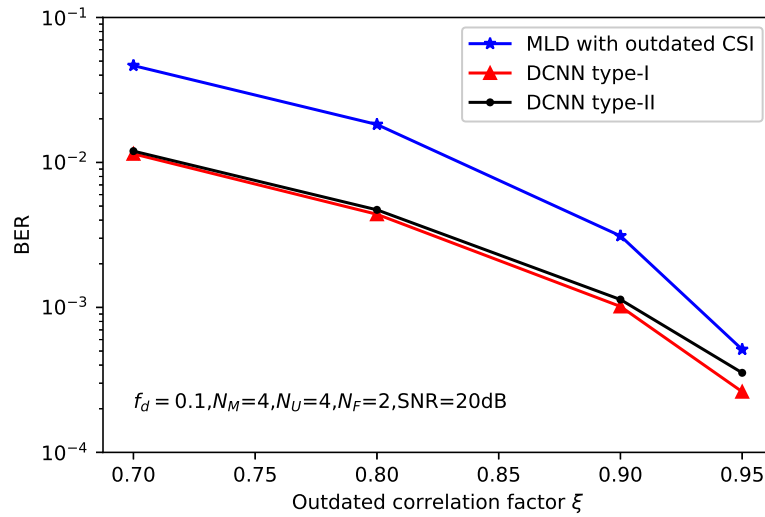
**Figure 8.** BER performance of deep learning-based MLD with SNR = 20 dB.

The effect of the normalized frequency $f_d$ is present in Figure 9, where the MIMO configuration remains the same with previous setup. The BER curves of both DCNN training models are provided with $f_d = 0.1$ and $f_d = 0.05$, respectively. We can see from this figure that smaller $f_d$ produces better performance with a gain of about 4dB. The reason is that if $f_d$ is smaller, the channel fading matrix changes more slowly, and the wireless channel can be learned more efficiently by the DCNN. As a result, more accurate CSI can be produced and enhancing the BER performance.
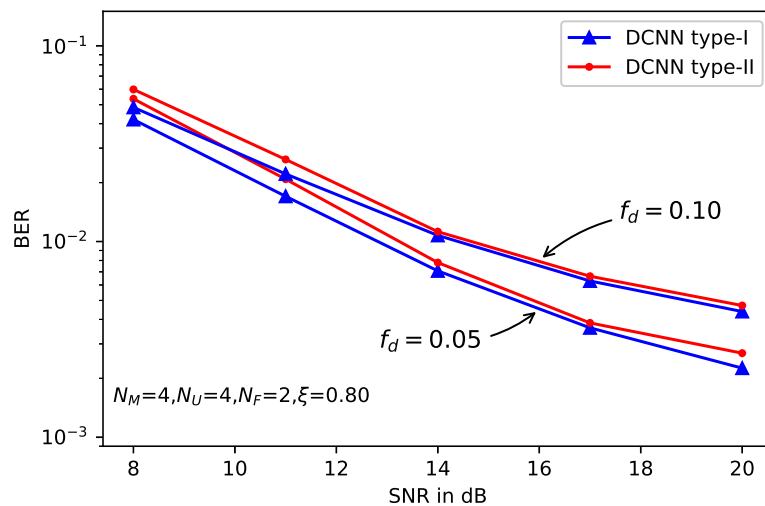


**Figure 9.** BER performance of deep learning-based MLD with different $f_d$.

Figures 10 and 11 depict the effects of the antenna number on BER performance with the normalized doppler frequency $f_d = 0.1$, and the correlation factor of imperfect CSI $\xi = 0.80$. Specifically, the antenna

configuration in Figure 10 is $N_M = 4, N_U = 4, N_F = 1$, and the number of data stream $N_D = 3$. In other words, the spectrum efficiency is higher than the previous setup. In Figure 11, the antenna configuration is $N_U = 2$ and $N_U = 3$, respectively. We can see from the two figures that the performance gain of DCNN with larger antenna number of user is obvious compared with the smaller antenna number, especially in the higher SNR region. Specifically, with $N_F = 1$, the performance gain of DCNN to the standard MLD is about 6dB. The performance gain of $N_U = 3$ is about 8dB than that of $N_U = 2$. That reason is that the larger receiver antennas' number can introduce more freedom of space diversity, thus the BER will be decreased greatly.
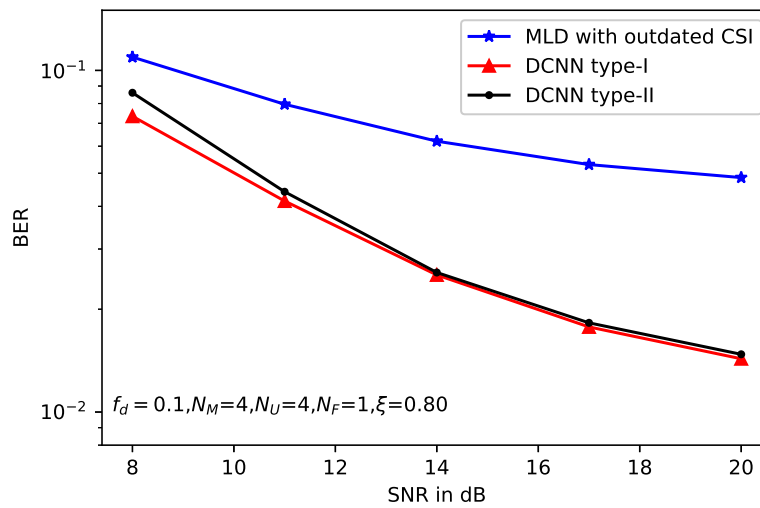


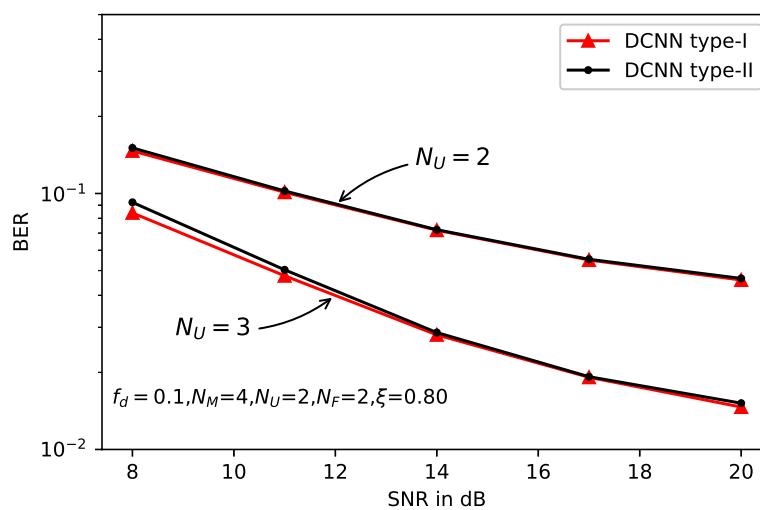**Figure 10.** BER performance of deep learning-based MLD with $N_F = 1$.



**Figure 11.** BER performance of deep learning-based MLD with different $N_U$.

## 6. Conclusions

In this paper, we investigate two types of deep learning-based secure MIMO detectors for heterogeneous networks. In the considered system, the equivalent channel fading matrix of the femto BS is zero-forced through null-space eigenvectors. The BER of the associated user is adopted as the metric to evaluate the system performance. with the help of deep convolutional neural networks, the macro BS produces more accurate CSI. The impacts of system parameters, such as the correlation factor of imperfect CSI, the normalized doppler frequency, the number of antennas are investigated in different setup scenarios. Considerable performance gains can be obtained from the deep learning-based detectors compared with the classical maximum likelihood algorithm.

## References

1. Wang, D.; Bai, B.; Lei, K.; Zhao, W.; Yang, Y.; Han, Z. Enhancing Information Security via Physical Layer Approaches in Heterogeneous IoT With Multiple Access Mobile Edge Computing in Smart City. *IEEE Access* **2019**, *7*, 54508–54521. doi:10.1109/ACCESS.2019.2913438. [CrossRef]

2. Li, X.; Li, J.; Liu, Y.; Ding, Z.; Nallanathan, A. Residual Transceiver Hardware Impairments on Cooperative NOMA Networks. *IEEE Trans. Inf. Forensics Secur.* **2020**, *19*, 680–695. [CrossRef]

3. Li, X.; Wang, Q.; Peng, H.; Zhang, H.; Do, D.-T.; Rabie, K.M.; Kharel, R.; CavalcanteA. A Unified Framework for HS-UAV NOMA Networks: Performance Analysis and Location Optimization. *IEEE Wirel. Commun. Lett.* **2020**, *8*, 13329–13340. doi:10.1109/ACCESS.2020.2964730. [CrossRef]

4. Dai, P.; Liu, K.; Wu, X.; Liao, Y.; Lee, V.C.S.; Son, S.H. Bandwidth Efficiency and Service Adaptiveness Oriented Data Dissemination in Heterogeneous Vehicular Networks. *IEEE Trans. Veh. Technol.* **2018**, *67*, 6585–6598. doi:10.1109/TVT.2018.2812742. [CrossRef]

5. Zhou, Y.; Yu, F.R.; Chen, J.; Kuo, Y. Resource Allocation for Information-Centric Virtualized Heterogeneous Networks With In-Network Caching and Mobile Edge Computing. *IEEE Trans. Veh. Technol.* **2017**, *66*, 11339–11351. doi:10.1109/TVT.2017.2737028. [CrossRef]

6. Zhong, Z.; Peng, J.; Huang, K.; Zhong, Z. Analysis on Physical-Layer Security for Internet of Things in Ultra Dense Heterogeneous Networks. In Proceedings of the 2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Chengdu, China, 15–18 December 2016; pp. 39–43. doi:10.1109/iThings-GreenCom-CPSCom-SmartData.2016.34. [CrossRef]

7. Wang, H.; Zheng, T.; Yuan, J.; Towsley, D.; Lee, M.H. Physical Layer Security in Heterogeneous Cellular Networks. *IEEE Trans. Commun.* **2016**, *64*, 1204–1219. doi:10.1109/TCOMM.2016.2519402. [CrossRef]

8. Wu, Y.; Khisti, A.; Xiao, C.; Caire, G.; Wong, K.; Gao, X. A Survey of Physical Layer Security Techniques for 5G Wireless Networks and Challenges Ahead. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 679–695. doi:10.1109/JSAC.2018.2825560. [CrossRef]

9. Tang, W.; Feng, S.; Ding, Y.; Liu, Y. Physical Layer Security in Heterogeneous Networks With Jammer Selection and Full-Duplex Users. *IEEE Trans. Wireless Commun.* **2017**, *16*, 7982–7995. doi:10.1109/TWC.2017.2755640. [CrossRef]

10. Ma, Z.; Lu, Y.; Shen, L.; Liu, Y.; Wang, N. Cooperative Jamming and Relay Beamforming Design for Physical Layer Secure Two-Way Relaying. In Proceedings of the 2018 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Zhengzhou, China, 18–20 October 2018; pp. 333–3336. doi:10.1109/CyberC.2018.00066. [CrossRef]

11. Yang, N.; Yan, S.; Yuan, J.; Malaney, R.; Subramanian, R.; Land, I. Artificial Noise: Transmission Optimization in Multi-Input Single-Output Wiretap Channels. *IEEE Trans. Commun.* **2015**, *63*, 1771–1783. doi:10.1109/TCOMM.2015.2419634. [CrossRef]

12. Zheng, T.; Wang, H.; Yuan, J.; Towsley, D.; Lee, M.H. Multi-Antenna Transmission With Artificial Noise Against Randomly Distributed Eavesdroppers. *IEEE Trans. Commun.* **2015**, *63*, 4347–4362. doi:10.1109/TCOMM.2015.2474390. [CrossRef]

13. Wang, W.; Teh, K.C.; Luo, S.; Li, K.H. Physical Layer Security in Heterogeneous Networks With Pilot Attack: A Stochastic Geometry Approach. *IEEE Trans. Commun.* **2018**, *66*, 6437–6449. doi:10.1109/TCOMM.2018.2859954. [CrossRef]

14. Zhao, W.; Chen, Z.; Li, K.; Liu, N.; Xia, B.; Luo, L. Caching-Aided Physical Layer Security in Wireless Cache-Enabled Heterogeneous Networks. *IEEE Access* **2018**, *6*, 68920–68931. doi:10.1109/ACCESS.2018.2880339. [CrossRef]

15. Zhu, J.; Gong, C.; Zhang, S.; Zhao, M.; Zhou, W. Foundation study on wireless big data: Concept, mining, learning and practices. *Chin. Commun.* **2018**, *15*, 1–15.

16. Deng, D.; Fan, L.; Zhao, R.; Hu, R.Q. Secure communications in multiple amplify-and-forward relay networks with outdated channel state information. *Trans. Emerging Telecommun. Technol.* **2016**, *27*, 494–503. [CrossRef]

17. Michalopoulos, D.S.; Suraweera, H.A.; Karagiannidis, G.K.; Schober, R. Amplify-and-Forward Relay Selection with Outdated Channel Estimates. *IEEE Trans. Commun.* **2012**, *60*, 1278–1290. [CrossRef]

18. Fan, L.; Lei, X.; Fan, P.; Hu, R. Outage probability analysis and power allocation for two-way relay networks with user selection and outdated channel state information. *IEEE Commun. Lett.* **2012**, *16*, 638–641. [CrossRef]

19. Li, X.; Li, J.; Li, L. Performance Analysis of Impaired SWIPT NOMA Relaying Networks Over Imperfect Weibull Channels. *IEEE Syst. J.* **2020**, *99*, 669–672. doi:10.1109/JSYST.2019.2919654. [CrossRef]

20. Li, X.; Liu, M.; Deng, C. Full-Duplex Cooperative NOMA Relaying Systems With I/Q Imbalance and Imperfect SIC. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 17–20. doi:10.1109/JSYST.2019.2919654. [CrossRef]

21. Li, X.; Li, J.; Li, L.; Jin, J.; Zhang, J.; Zhang, D. Effective Rate of MISO Systems Over $\kappa$ - $\mu$ Shadowed Fading Channels. *IEEE Access* **2017**, *5*, 10605–10611. [CrossRef]

22. Wu, Y.; Louie, R.H.Y.; McKay, M.R. Analysis and Design of Wireless Ad Hoc Networks With Channel Estimation Errors. *IEEE Trans. Signal Process.* **2013**, *61*, 1447–1459. [CrossRef]

23. Savazzi, S.; Spagnolini, U. Optimizing Training Lengths and Training Intervals in Time-Varying Fading Channels. *IEEE Trans. Signal Process.* **2009**, *57*, 1098–1112. [CrossRef]

24. Han, S.; Tian, Y.; Yang, C. User-Specified Training Symbol Placement for Channel Prediction in TDD MIMO Systems. *IEEE Trans. Veh. Technol.* **2011**, *60*, 2837–2843. [CrossRef]

25. Soltani, M.; Pourahmadi, V.; Mirzaei, A.; Sheikhzadeh, H. Deep Learning-Based Channel Estimation. *IEEE Commun. Lett.* **2019**, *23*, 652–655. doi:10.1109/LCOMM.2019.2898944. [CrossRef]

26. Balevi, E.; Andrews, J.G. One-Bit OFDM Receivers via Deep Learning. *IEEE Trans. Commun.* **2019**, *67*, 4326–4336. doi:10.1109/TCOMM.2019.2903811. [CrossRef]

27. Wu, Y.; Li, X.; Fang, J. A Deep Learning Approach for Modulation Recognition via Exploiting Temporal Correlations. In Proceedings of the 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), kalamata, Greece, 25–28 June 2018; pp. 1–5. doi:10.1109/SPAWC.2018.8445938. [CrossRef]

28. Yuan, J.; Ngo, H.Q.; Matthaiou, M. Machine Learning-Based Channel Prediction in Massive MIMO with Channel Aging. *IEEE Trans. Wireless Commun.* **2020**, 1. doi:10.1109/TWC.2020.2969627. [CrossRef]

29. Amirabadi, M.A. Deep learning for channel estimation in FSO communication system. *arXiv* **2020**, arXiv:1909.11003. doi:10.1016/j.optcom.2019.124989.

30. Amirabadi, M.A. A deep learning based solution for imperfect CSI problem in correlated FSO communication channel. *arXiv* **2020**, arXiv:1909.11002.

31. Yang, N.; Elkashlan, M.; Duong, T.Q.; Yuan, J.; Malaney, R. Optimal Transmission With Artificial Noise in MISOME Wiretap Channels. *IEEE Trans. Veh. Technol.* **2016**, *65*, 2170–2181. doi:10.1109/TVT.2015.2419318. [CrossRef]

32. Fritschek, R.; Schaefer, R.F.; Wunder, G. Deep Learning for the Gaussian Wiretap Channel. In Proceedings of the 2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–6. doi:10.1109/ICC.2019.8761681. [CrossRef]

33. Xing, J.; Lv, T.; Zhang, X. Cooperative Relay Based on Machine Learning for Enhancing Physical Layer Security. In Proceedings of the 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Istanbul, Turkey, 8–11 September 2019; pp. 1–6. doi:10.1109/PIMRC.2019.8904319. [CrossRef]

34. Xiao, L.; Sheng, G.; Liu, S.; Dai, H.; Peng, M.; Song, J. Deep Reinforcement Learning-Enabled Secure Visible Light Communication Against Eavesdropping. *IEEE Trans. Commun.* **2019**, *67*, 6994–7005. doi:10.1109/TCOMM.2019.2930247. [CrossRef]

35. He, D.; Liu, C.; Wang, H.; Quek, T.Q.S. Learning-Based Wireless Powered Secure Transmission. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 600–603. doi:10.1109/LWC.2018.2881976. [CrossRef]

36. Xiao, L.; Li, Y.; Han, G.; Dai, H.; Poor, H.V. A Secure Mobile Crowdsensing Game With Deep Reinforcement Learning. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 35–47. doi:10.1109/TIFS.2017.2737968. [CrossRef]

37. Simon, M.K.; Alouini, M.S. *Digital Communication Over Fading*, 2nd ed; Wiley: Hoboken, NJ, USA, 2005.

38. Ji, Y.; Fan, W.; Kyösti, P.; Li, J.; Pedersen, G.F. Antenna Correlation Under Geometry-Based Stochastic Channel Models. *IEEE Antennas Wirel. Propag. Lett.* **2019**, *18*, 2567–2571. doi:10.1109/LAWP.2019.2943413. [CrossRef]

39. Jiang, Z.; Chen, S.; Molisch, A.F.; Vannithamby, R.; Zhou, S.; Niu, Z. Exploiting Wireless Channel State Information Structures Beyond Linear Correlations: A Deep Learning Approach. *IEEE Commun. Mag.* **2019**, *57*, 28–34. doi:10.1109/MCOM.2019.1800581. [CrossRef]

40. O'Shea, T.J.; Corgan, J.; Clancy, T.C. Convolutional Radio Modulation Recognition Networks. In *Engineering Applications of Neural Networks*; Jayne, C.; Iliadis, L., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 213–226.

41. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.

42. He, K.; Zhang, X.; Ren, S.; Jian, S. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016.

43. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

44. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. Available online: http://download.tensorflow.org/paper/whitepaper2015.pdf (accessed on 20 March 2020)