

Article

# Deep-Learning-Based Frame Format Detection for IEEE 802.11 Wireless Local Area Networks

Minjae Kim, Zhongfeng Zhang, Daejin Kim and Seungwon Choi \*

Department of Electronic and Computer Engineering, Hanyang University, Wangsimni-ro 2222, Seongdong-gu, Seoul 04763, Korea; minjae.kim@dsplab.hanyang.ac.kr (M.K.); zhongfeng.zhang@dsplab.hanyang.ac.kr (Z.Z.); kdj317@dsplab.hanyang.ac.kr (D.K.)

\* Correspondence: choi@dsplab.hanyang.ac.kr

Received: 22 June 2020; Accepted: 16 July 2020; Published: 18 July 2020



**Abstract:** Backward compatibility is one of the key issues for radio equipment that supports IEEE 802.11, which is a typical communication protocol for wireless local area networks (WLANs). For achieving successful packet decoding with backward compatibility, frame format detection is the core precondition. In this study, we present a novel, deep-learning-based frame format detection method for IEEE 802.11 WLANs. Considering that the detection performance of conventional methods is mainly degraded because of poor performance in symbol synchronization and/or channel estimation in environments with a low signal-to-noise ratio, we propose a novel detection method based on a deep learning network to replace conventional detection procedures. The proposed deep-learning network method achieves robust detection directly from the received (Rx) data. Through extensive computer simulations performed in multipath fading channel environments (modeled by Project IEEE 802.11 Task Group ac), we confirmed that the proposed method exhibits significantly higher frame format detection performance than that of the conventional method.

**Keywords:** deep learning network; wireless local area network; IEEE 802.11ac; backward compatibility; frame format detection

## 1. Introduction

As the new standards of wireless local area networks (WLANs) affiliated with IEEE 802.11 [1] are being released for supporting tremendously increasing data traffic, the problem of backward compatibility, which requires a single piece of WLAN radio equipment to support the entire IEEE 802.11 family, has become more critical than ever before [2]. For instance, radio equipment for IEEE 802.11ac [3] is also required to support the previous standards, i.e., IEEE 802.11a/n.

The first step to support the aforementioned backward compatibility is to correctly detect the frame format of each standard such that the receiver can distinguish different standards from one another. This implies that radio equipment for IEEE 802.11ac should support not only the frame format of very-high throughput (VHT) (i.e., developed for IEEE 802.11ac), but also the frame formats of non-high throughput (non-HT) and high throughput mixed format (HT-MF), both of which were originally developed for IEEE 802.11a and IEEE 802.11n, respectively.

Therefore, performing robust frame format detection is crucial, especially in an adverse signal environment with a significantly low signal-to-noise ratio (SNR). However, despite the importance of frame format detection, the conventional frame format detection method [4], whose performance substantially depends upon symbol synchronization and channel compensation, is significantly vulnerable to various impediments, such as a low SNR and/or multipath fading of a given signal environment. Furthermore, the interference problem in the unlicensed bands of 2.4 or 5 GHz, which IEEE 802.11 uses, would make the conventional method of frame format detection even more

unreliable, as the number of Internet of Things devices and long-term evolution unlicensed devices occupying the same unlicensed band has drastically increased recently.

Because deep-learning-based methods have been widely used in various research fields, including computer science and natural language processing, the concept of deep learning might be applied to mobile communication networks including WLANs, especially for the problem of frame format detection. The initial stage of applying deep learning in networking-related areas was mainly performed for the upper layer communication protocols [5]. Recently, several meaningful studies on applying deep learning to communications focused on physical layer signal processing, including signal detection, modulation classification, and channel decoding [6,7].

In this study, we present a novel deep-learning-based procedure for detecting the frame formats of various versions of IEEE 802.11. We aim to detect the frame format of the received (Rx) data, such that the receiver can support all of the different versions of IEEE 802.11 on the basis of the detected frame format prior to decoding the Rx data. The most important motivation behind adopting deep learning is that it is significantly more robust than conventional methods, especially when the channel falls into a deep fade such that the resultant SNR becomes so low that the conventional method cannot perform channel estimation (CE) and channel compensation to correctly detect the frame format [4].

Deep learning is considered to be significantly better than the conventional method, especially in a low SNR situation, because it is considerably more robust for classification problems in many other applications, including various types of image classifications [8]. As shown in subsequent sections, the performance of the deep-learning-based frame format detection procedure is significantly more robust than those of the conventional methods based on symbol synchronization and channel compensation. In this study, the Rx data are appended to the proposed deep learning network such that the proposed system outputs the frame detection results without explicitly performing synchronization, CE, and compensation, thereby achieving an excellent performance.

The remainder of this paper is organized in the following structure. In Section 2, the detection problem is summarized with the introductions of both the conventional and proposed deep learning methods. In Section 3, the simulations performed in six different fading channel environments, modeled by Project IEEE 802 Task Group ac, are described and the results presented. In Section 4, the conclusions drawn from the simulation results are summarized.

## 2. Detection Method

We first discuss a conventional method for detecting the frame format of IEEE 802.11 Rx data. Considering the aforementioned intrinsic problems of the conventional method, we propose a novel, seven-layer, deep-learning-based frame format detection procedure. The objective is to identify which of the three frame formats of IEEE 802.11 a/n/ac, i.e., non-HT, HT-MF, and VHT, respectively, corresponds to that of the Rx data.

### 2.1. Conventional Frame Format Detection Method

Figure 1 illustrates three frame formats used in IEEE 802.11a/n/ac: non-HT, HT-MF, and VHT. Because the three symbols after the legacy-long training field (L-LTF) adopt binary phase shift keying (BPSK) and quadrature BPSK (QBPSK) with different combinations, as shown in Figure 1, they can be used to distinguish the frame formats from one another.

Figure 2 illustrates the constellation diagrams of BPSK and QBPSK, respectively, employed in the IEEE 802.11 family shown in Figure 1.

The conventional method for detecting the frame format of Rx data is primarily based on symbol synchronization and CE. Particularly, we consider a typical frame format detection procedure for IEEE 802.11 WLANs, as shown in Figure 3. Before performing frame format detection, the Rx data should be processed via symbol synchronization, L-LTF extraction, L-LTF demodulation, and CE/channel compensation, as shown in Figure 3. Subsequently, the three symbols after the aforementioned L-LTF are used for frame format detection.

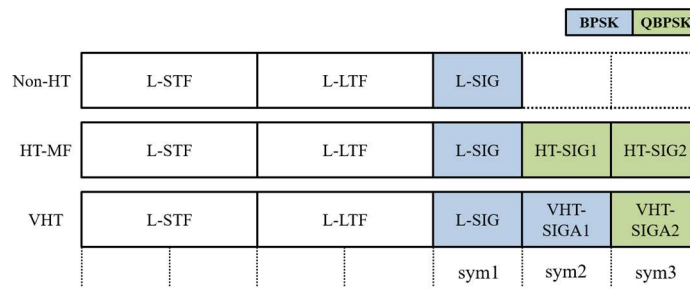


Figure 1. The three frame formats adopted in IEEE 802.11a/n/ac.

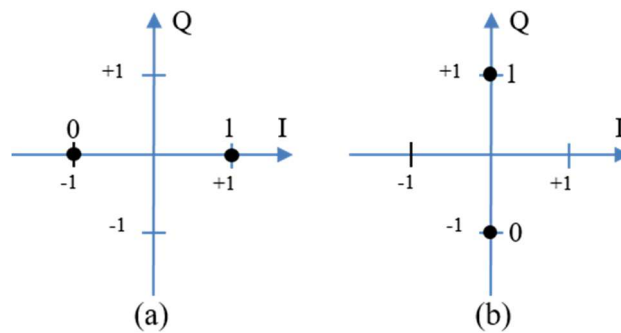


Figure 2. Constellation diagrams of (a) binary phase shift keying (BPSK), and (b) quadrature BPSK (QBPSK).

First, because the BPSK coding rate of non-HT can be 1/2 or 3/4, whereas those of HT-MF and VHT can only be 1/2, the modulation coding scheme of the first symbol (among the three) determines whether the frame format is non-HT. Subsequently, if the second symbol (among the three after the L-LTF) is modulated according to QBPSK, the frame format is detected as HT-MF; otherwise, the detection procedure should continue. Finally, if the third symbol is modulated via QBPSK, the frame format is VHT, and otherwise it is classified as non-HT.

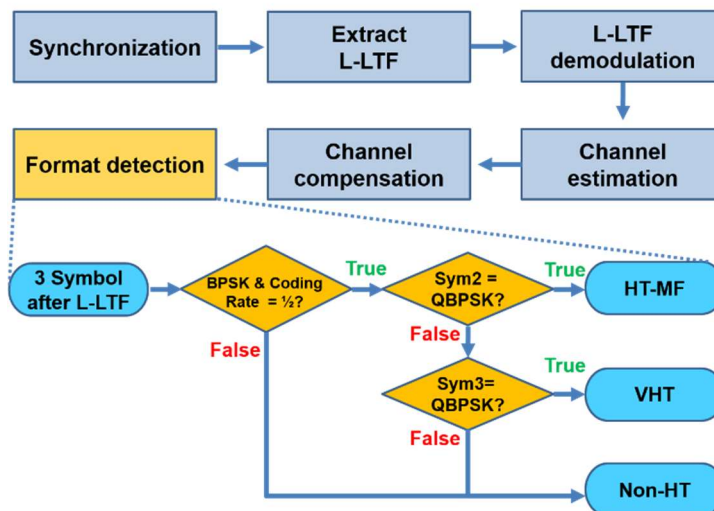


Figure 3. Flowchart of the conventional frame format detection procedure.

2.2. Deep-Learning-Based Method

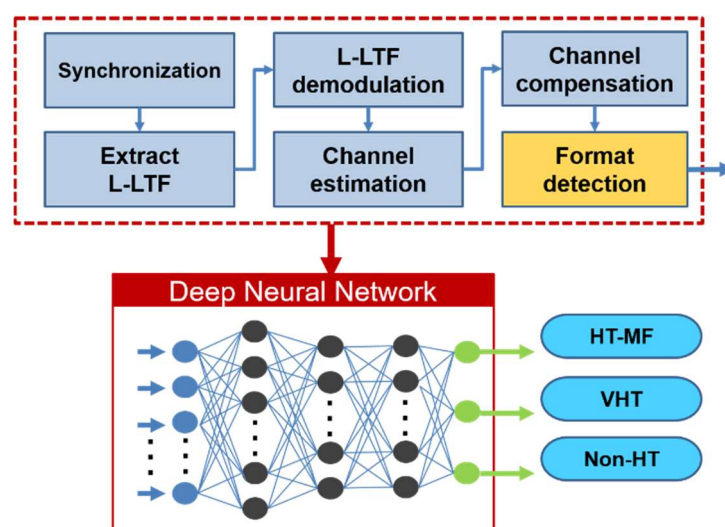
The deep neural network (DNN) model considered in this study comprises seven layers, as shown in Table 1. The chosen hyperparameters, such as the number of layers and neurons, are decided by

conducting multiple experiments on different combinations of hyperparameters and proved to have the best performance among them. The input layer takes eight symbols as its inputs: four symbols after the L-LTF (i.e., sym1, sym2, and sym3, as shown in Figure 1, and one more data field symbol in addition to two symbols in the legacy-short training field (L-STF) and two symbols in the L-LTF. Notably, all four symbols, both in the L-STF and L-LTF, are known as training symbols. Particularly, this implies that the input layer takes as its inputs 640 complex-valued samples (= 80 samples/symbol), resulting in 640 input components for the real and imaginary parts, respectively. In the case of hidden layers, rather than opting for the commonly used rectified linear unit (ReLU), we adopted a scaled exponential linear unit (SELU) [9] as the activation function to express negative values, thereby avoiding the dying ReLU phenomenon [10]. The softmax function has been adopted as the activation function for the output layer because the sum of the output (i.e., the entire probability), should be 1. Meanwhile, categorical cross entropy has been used as the loss function because our problem is to classify multiple objects rather than binary objects. Furthermore, we have adopted the adaptive moment estimation (Adam) optimizer, whose step size is obtained from the gradient value of the previous step [11].

**Table 1.** Configuration of the proposed seven-layer deep learning network.

Input	1280 (640 Complex Samples)
Dense (SELU)	256
Dense (SELU)	128
Dense (SELU)	64
Dense (SELU)	32
Dense (SELU)	16
Output (Softmax)	3
Initial weight value	He_uniform
Loss function	Categorical cross entropy
Optimizer	Adam
Learning rate	0.001
Batch size	128

As previously explained, we have formed a DNN (see Figure 4) that resolves all of the procedures, including symbol synchronization, CE, and channel compensation, required in the conventional procedure into one DNN-based black-box.



**Figure 4.** Block diagram of the proposed deep-learning-based frame format detection method.

### 3. Simulations and Numerical Results

This section first presents the generation of transmitted (Tx) data and parameter setup of the Rx data for computer simulations. Subsequently, the simulation results obtained using the proposed deep learning procedure shown in Figure 4 are compared with those obtained using the conventional method (which is based on CE) shown in Figure 3.

As shown in Figure 5, the Tx and Rx data used in our simulations were generated using the MATLAB WLAN toolbox [12], and the Rx data parameter values are presented in Table 2.

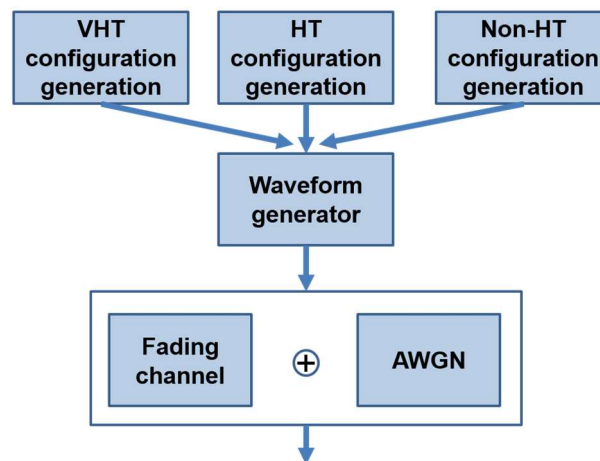


Figure 5. Transmitted/received (Tx/Rx) data generation.

As the carrier frequency, IEEE 802.11a uses 5 GHz, IEEE 802.11n uses 2.4 and 5 GHz, and IEEE802.11ac uses 5 GHz [3]. As the signal bandwidth, non-HT uses 5, 10, or 20 MHz, while HT employs 20 or 40 MHz. In the case of VHT, a bandwidth of 20, 40, 80, or 160 MHz is used. To perform simulations (i.e., to detect the frame format), we set the carrier frequency and bandwidth to be 5 GHz and 20 MHz, respectively, which maximizes the vagueness among the three frame formats. Additionally, to not allow the frame format detection to be easily noticed simply from the modulation difference in the data field, we adopted BPSK for the entire data field of all the three candidates, i.e., IEEE 802.11 a/n/ac. Because the first symbol after the L-LTF in all of the three candidates employs BPSK, the vagueness among the three frame formats would be further enhanced.

Table 2. Rx-data parameters.

Class	VHT, HT-MF, and Non-HT
Carrier frequency	5 GHz
Bandwidth	20 MHz
Channel	IEEE 802.11ac multipath fading channel (Models A–F)
Noise	AWGN
SNR	– 10 to 20 dB
Data field modulation scheme	BPSK

For modeling the channel, we adopted the wireless, fading channel modeling suggested by Project IEEE 802.11 Task Group ac, which provided six different models (i.e., Models A–F), as shown in Table 2 [13,14]. In the modeled wireless fading channel, the Rx data of non-HT, HT-MF, and VHT formats were generated for each modeled channel with the SNR being randomly varied from –10 to 20 dB with a uniform distribution. The configuration of the training data set was assigned to each of the six channel models, i.e., Models A–F. The total number of training data sets generated for each of the three formats was 500 per SNR value of 1 dB intervals. Because we considered SNR values in the range of –10 to 20 dB, the total number of training data sets

generated for the detection procedure is given as follows: 500 (dataset/format) × 31 (number of observed SNR values) × 3 (formats) × 8 (symbols/dataset) × 80 (samples/symbol) × 6 (channel models). Notably, each data set comprises eight symbols (i.e., two symbols in the L-STF, two in the L-LTF, and four after the L-LTF), suggesting that we were considering one more symbol after the “sym3” shown in Figure 1. The reason for the one extra symbol so that the total number of samples became eight instead of seven was to tackle the delay (timing offset) caused by the multipath fading channel.

As for the testing data, we generated twice the amount of training data by using the same procedure as that used to generate the training data. The correct detection probability worsened as the channel shifted from Model-A to Model-F. Table 3 summarizes the channel modeling provided by the Project IEEE 802.11 Task Group ac. Notably, Model-A corresponds to a residential environment, while Model-F corresponds to an urban environment. Figure 6 illustrates the detection performances of the proposed deep-learning-based method and the CE-based conventional method. It also presents the performance when the symbol synchronization is not included in the deep learning network for comparison purposes. As shown in the figure, the conventional method provided approximately 90% accuracy when the SNR was 5 dB. Comparatively, the proposed deep-learning-based method exhibited a comparable performance when the SNR was −10 dB, resulting in approximately 15 dB of SNR improvement. The conventional method required approximately 15 dB of SNR for achieving 99% accuracy, which was achieved by the proposed method with only 0 dB of SNR. Furthermore, when symbol synchronization was not included in the deep learning network, the performance improvement was limited to approximately 5 dB, as shown in the figure. We can see that the deep-learning-based method showed excellent performance in all of the channel models with negligible differences. It significantly lowered the complexity of the system for not having to train and apply different models to different channel models accordingly.

Table 3. IEEE 802.11 task group ac multipath fading channel.

Parameter	Model					
	A	B	C	D	E	F
Breakpoint distance (m)	5	5	5	10	20	30
RMS delay spread (ns)	0	15	30	50	100	150
Maximum delay (ns)	0	80	200	390	730	1050
Rician K-factor (dB)	0	0	0	3	6	6
Number of taps	1	9	14	18	18	18
Number of clusters	1	2	2	3	4	6

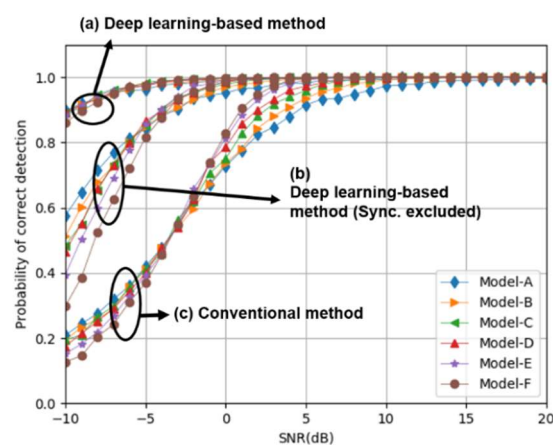


Figure 6. Format detection performances: (a) proposed deep learning method, (b) deep learning with symbol synchronization provided by the conventional method, and (c) conventional method based on symbol synchronization and channel estimation (CE)/channel compensation.

Figure 7 illustrates the confusion matrices provided by the proposed deep-learning-based method when the channel model is fixed to Model-F, which was the most adverse signal environment among the six modeled channels. As shown in the figure, when the channel model was fixed to Model-F, the frame format detection accuracy became higher as the SNR increased. When the SNR was 0 dB, the accuracy for VHT, HT-MF, and non-HT format detection reached almost 100%. From our simulations, it is obvious that the deep-learning-based method provides significantly better performance than that of the conventional method, for the frame format detection in IEEE 802.11 WLANs.

As for the overhead of using the deep-learning-based method, we measured the processing time during the testing stage. It turned out that the processing time on average was as short as 0.1 ms. Note that the processing time measured in the simulations was largely limited to the hardware that we used, and could have been significantly shortened by employing a more advanced graphic processing unit (GPU), which could provide more powerful parallel computing capabilities. Therefore, we believe that it is promising to apply the deep-learning-based method for format detection to the practical signal environment of WLAN systems.

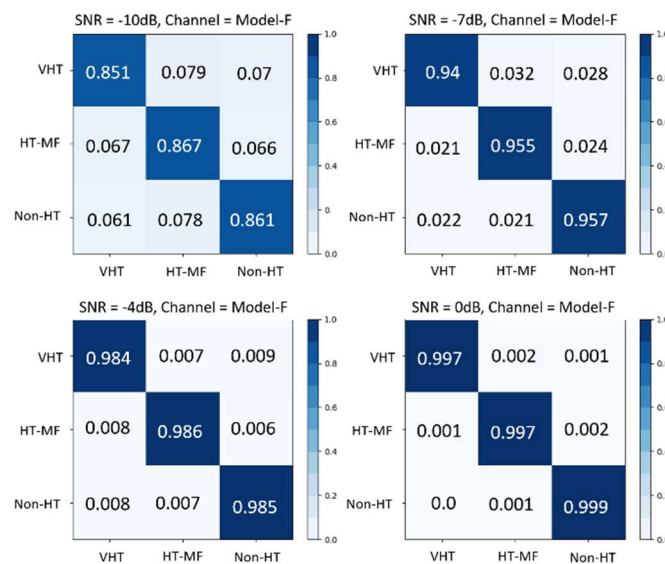
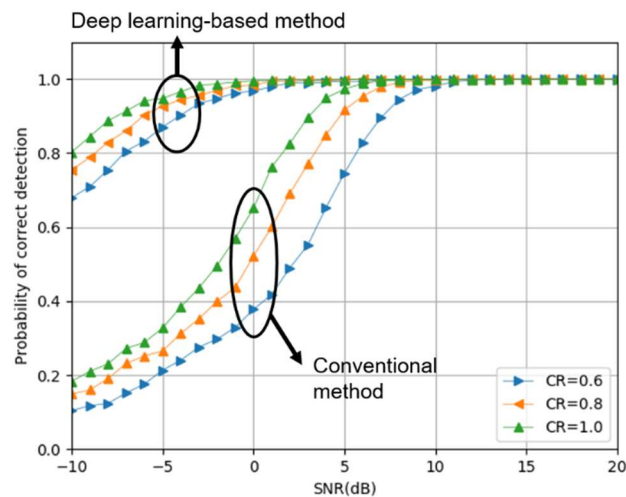


Figure 7. Confusion matrix for different signal-to-noise ratios (SNRs) in Model-F.

Next, the proposed deep-learning-based method was applied to another signal environment wherein the Tx data set included clipping noise. To reduce the peak-to-average power ratio in orthogonal frequency division multiplexing (OFDM) systems, clipping is often included in the Tx data generated [15], which causes additional nonlinear noises in the Rx data.

Figure 8 compares the frame format detection performances of the proposed deep-learning-based method with those of the CE-based conventional method when the signal environment is fixed to Model-F, which was the most adverse channel among the six channel models. In our simulations, the clipping ratio (CR) was arbitrarily set to 1, 0.8, and 0.6. As shown in the figure, as the CR decreased, which implied that the Tx signal was clipped more severely, the detection accuracy deteriorated in general. However, the proposed deep-learning-based method exhibited significantly more robust performance against clipping than that of the conventional one.



**Figure 8.** Frame format detection accuracy of deep-learning-based and conventional methods according to three clipping ratio (CR) values.

#### 4. Conclusions

We adopted a seven-layer deep learning network for the frame format detection of IEEE 802.11 WLANs. Through simulations, we observed that the proposed deep-learning-based detection procedure achieved excellent performance compared with the conventional method, especially when the SNR of the wireless channel was so low that the conventional method failed to perform symbol synchronization and/or channel estimation. We also observed that the deep-learning-based format detection exhibited a significant performance improvement compared with the conventional method when signal clipping happened due to the limitations of amplifiers. In addition to the significant performance improvement, the deep-learning-based method also proved to behave consistently with excellent performances for all of the channel models in the simulations.

In this paper, we did not address issues such as the system complexity when implementing deep-learning-based IEEE 802.11 format detection systems since this paper is mainly focused on revealing the potential usage of deep learning in WLAN systems, as well as its significant performance improvement compared with the conventional method. However, the implementation of WLAN systems based on deep learning will be our future work.

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

**Acknowledgments:** This work was supported by an Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (No. 2017-0-00723, Development of Software-defined Service-oriented Integrated base station platform using Reconfigurable Radio System Technology).

**Conflicts of Interest:** The authors declare no conflicts of interest.

#### References

1. IEEE 802.11 WLANs, Working Group for WLAN Standards. Available online: <https://www.iol.unh.edu/knowledge/ieee-80211-working-group-wlan-standards> (accessed on 22 June 2020).
2. Lopez-Aguilera, E.; Garcia-Villegas, E.; Casademont, J. Evaluation of IEEE 802.11 coexistence in WLAN deployments. *Wireless. Netw.* **2019**, *25*, 87–104. [CrossRef]
3. Perahia, E.; Stacey, R. *Next Generation Wireless LANs: 802.11n and 802.11ac*; Cambridge University Press: Cambridge, UK, 2013.

4. Chanvali, N.K. Maximum likelihood detection of a frame format in HT and VHT wireless LANs. In Proceedings of the International Conference on Electronics, Computing and Communication Technologies (IEEE CONECCT), Bangalore, India, 17–19 January 2013.
5. Zhang, C.; Patras, P.; Haddadi, H. Deep learning in mobile and wireless networking: A survey. *IEEE Commun. Surveys Tutorials*. **2019**, *21*, 2224–2287. [CrossRef]
6. O’Shea, T.; Hoydis, J. An introduction to deep learning for the physical layer. *IEEE Trans. Cogn. Commun. Netw.* **2017**, *3*, 563–575. [CrossRef]
7. Ye, H.; Li, G.Y.; Juang, B.-H.F. Power of deep learning for channel estimation and signal detection in OFDM systems. *IEEE Wireless Commun. Lett.* **2017**, *7*, 114–117.
8. Triantaris, P.; Tsimbaló, E.; Chin, W.H.; Gündüz, D. Automatic modulation classification in the presence of interference. Proceedings of European Conference on Networks Communication, Valencia, Spain, 18–21 June 2019; pp. 549–553.
9. Zhang, G.; Li, H. Effectiveness of scaled exponentially-regularized linear units (SERLUs). Available online: <https://arxiv.org/abs/1807.10117> (accessed on 22 June 2020).
10. Lu, L.; Shin, Y.; Su, Y.; Karniadakis, G.E. Dying ReLU and initialization: Theory and numerical examples. Available online: <https://arxiv.org/abs/1903.06733> (accessed on 22 June 2020).
11. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. Available online: <https://arxiv.org/abs/1412.6980> (accessed on 22 June 2020).
12. *MATLAB and WLAN Toolbox Release 2018b*; The MathWorks, Inc.: Natick, MA, USA, 2018.
13. TGN Channel Models, IEEE Std. 802.11 -03/940r4. Available online: <https://mentor.ieee.org/802.11/dcn/03/11-03-0940-04-000n-tgn-channel-models.doc> (accessed on 22 June 2020).
14. TGac Channel Model Addendum Supporting Material, IEEE Std. 802.11-09/0569r0. Available online: <https://mentor.ieee.org/802.11/dcn/09/11-09-0308-12-00ac-tgac-channel-model-addendum-document.doc> (accessed on 22 June 2020).
15. Li, X.D.; Cimini Jr., L.J. Effects of clipping and filtering on the performance of OFDM. *IEEE Commun. Lett.* **1998**, *2*, 131–133.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).