

# Deep Learning at the Physical Layer: System Challenges and Applications to 5G and Beyond

Francesco Restuccia and Tommaso Melodia

The authors discuss the need for real-time DL at the physical layer, and then summarize the current state of the art and existing limitations. They conclude the article by discussing an agenda of research challenges and how DL can be applied to address crucial problems in 5G and beyond networks.

## ABSTRACT

The unprecedented requirements of IoT have made fine-grained optimization of spectrum resources an urgent necessity. Thus, designing techniques able to extract knowledge from the spectrum in real time and select the optimal spectrum access strategy accordingly has become more important than ever. Moreover, 5G networks will require complex management schemes to deal with problems such as adaptive beam management and rate selection. Although deep learning (DL) has been successful in modeling complex phenomena, commercially available wireless devices are still very far from actually adopting learning-based techniques to optimize their spectrum usage. In this article, we first discuss the need for real-time DL at the physical layer, and then summarize the current state of the art and existing limitations. We conclude the article by discussing an agenda of research challenges and how DL can be applied to address crucial problems in 5G and beyond networks.

## INTRODUCTION

The wireless spectrum is undeniably one of nature's most complex phenomena. This is especially true in the highly dynamic context of the Internet of Things (IoT), where the widespread presence of tiny embedded wireless devices seamlessly connected to people and objects will make spectrum-related quantities such as fading, noise, interference, and traffic patterns hardly predictable with traditional mathematical models. Techniques able to perform real-time fine-grained spectrum optimization will thus become fundamental to squeeze out any spectrum resource available to wireless devices.

There are a number of key issues — summarized in Fig. 1 — that make existing wireless optimization approaches not completely suitable to address the spectrum challenges mentioned above. On one hand, model-driven approaches aim to mathematically formalize the entirety of the network and optimize an objective function. Although yielding optimal solutions, these approaches are usually NP-hard, and thus unable to be run in real time. Moreover, they rely on a series of modeling assumptions (fading/noise distribution, traffic and mobility patterns, etc.) that

may not always be valid. On the other hand, *protocol-driven* approaches consider a specific wireless technology (e.g., WiFi, Bluetooth, or Zigbee) and attempt to heuristically change parameters including modulation scheme, coding level, packet size, and so on, based on metrics computed in real time from pilots and/or training symbols. Protocol-driven approaches, being heuristic in nature, necessarily yield sub-optimal performance.

To obtain the best of both worlds, a new approach called *spectrum-driven* is being explored. In short, by using real-time machine learning (ML) techniques implemented in the hardware portion of the wireless platform, we can design wireless systems that can learn by themselves the optimal spectrum actions to take given the current spectrum state. Concretely speaking, the big picture is to realize systems able to distinguish on their own different spectrum states (e.g., based on noise, interference, channel occupation, and similar), and change their hardware and software fabric in real time to implement the optimal spectrum action [1, 2]. However, despite the numerous recent advances, so far truly self-adaptive and self-resilient cognitive radios have been elusive. On the other hand, the success of deep learning at the physical layer (PHY-DL) in addressing problems such as modulation recognition [3], radio fingerprinting [4], and medium access control (MAC) [5] has taken us many steps in the right direction [6]. Thanks to its unique advantages, deep learning (DL) can really be a game changer, especially when cast in the context of a real-time hardware-based implementation.

Existing work has mostly focused on generating spectrum data and training models in the cloud. However, a number of key *system-level issues* still remain substantially unexplored. To this end, we notice that the most relevant survey work [7, 8] introduces research challenges from a learning perspective only. Moreover, [9] and similar survey work focuses on the application of DL to upper layers of the network stack. Since DL was not conceived having the constraints and requirements of wireless communications in mind, it is still unclear what the fundamental limitations of PHY-DL are. Moreover, existing work has still not explored how PHY-DL can be used to address problems in the context of 5G and beyond (5GB) networks. For this reason, the first key contribu-

tion of this article is to discuss the research challenges of real-time PHY-DL without considering any particular frequency band or radio technology. The second key contribution is the introduction of a series of practical problems that may be addressed by using PHY-DL techniques in the context of 5G. Notice that 5G networks are expected to be heavily based on millimeter-wave (mmWave) and ultra-wideband communications, hence our focus on these issues. Since an exhaustive compendium of the existing work in PHY-DL is outside the scope of this manuscript, we refer the reader to [9] for an excellent survey.

## WHY DEEP LEARNING AT THE PHYSICAL LAYER?

DL is exceptionally suited to address problems where closed-form mathematical expressions are difficult to obtain [10]. For this reason, convolutional neural networks (CNNs) are now being “borrowed” by wireless researchers to address handover and power management in cellular networks, dynamic spectrum access, resource allocation/slicing/caching, video streaming, and rate adaptation, just to name a few. Figure 2 summarizes why traditional ML may not effectively address real-time physical-layer problems. Overall, DL relieves from the burden of finding the right “features” characterizing a given wireless phenomenon. At the physical layer, this key advantage becomes almost a necessity for at least three reasons, which are discussed below.

### HIGHLY DIMENSIONAL FEATURE SPACES

Classifying waveforms ultimately boils down to distinguishing small-scale patterns in the in-phase-quadrature (I/Q) plane, which may not be clearly separable in a low-dimensional feature space. For example, in radio fingerprinting we want to distinguish among hundreds (potentially thousands) of devices based on the unique imperfections imposed by the hardware circuitry. While legacy low-dimensional techniques can correctly distinguish up to a few dozen devices [11], DL-based classifiers can scale up to hundreds of devices by learning extremely complex features in the I/Q space [4]. Similarly, O’Shea *et al.* [3] have demonstrated that on the 24-modulation dataset considered, DL models achieve on the average about 20 percent higher classification accuracy than legacy learning models under noisy channel conditions.

### ALL-IN-ONE APPROACH

The second key advantage of DL is that automatic feature extraction allows the system designer to reuse the same DL architecture — and thus, the same hardware circuit — to address different learning problems. This is because, as we explain below, CNNs learn I/Q patterns in the I/Q plane, making them amenable to address different classification problems. Existing work, indeed, has demonstrated that CNNs can be used for very different problems, ranging from modulation recognition [3] to radio fingerprinting [4]. CNNs also keep latency and energy consumption constant, as explained above. Figure 2 shows an example where a learning system is trained to classify modulation, number of carriers, and fingerprinting.

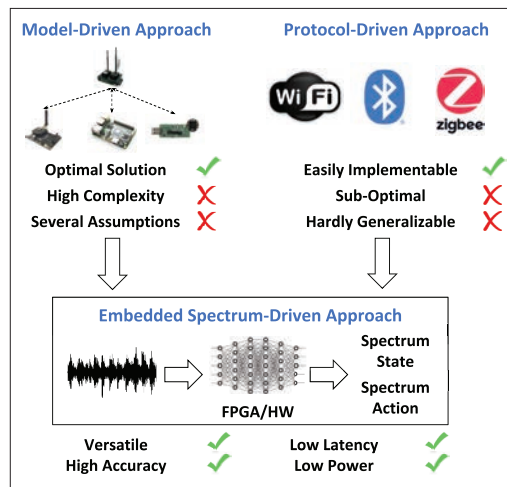


Figure 1. Key issues in today’s wireless optimization approaches.

While DL can concurrently recognize the three parameters, traditional learning requires different feature extraction processes for each of the classification outputs. This, in turn, increases hardware consumption and hinders fine-tuning of the learning model.

### REAL-TIME FINE TUNING

Model-driven optimization offers predictable performance only when the model actually matches the reality of the underlying phenomenon being captured. This implies that model-driven systems can yield sub-optimal performance when the model assumptions are different from what the network is actually experiencing. For example, a model assuming a Rayleigh fading channel can yield incorrect solutions when placed in a Rician fading environment. By using a data-driven approach, PHY-DL may easily be fine-tuned through the usage of fresh spectrum data, which can be used to find in real time a better set of parameters through gradient descent. Hand-tailored model-driven systems may result in being hard to fine-tune, as they might depend on a set of parameters that are not easily adaptable in real time (e.g., channel model). While DL “easily” accomplishes this goal by performing batch gradient descent on fresh input data, the same is not true for traditional ML, where tuning can be extremely challenging since it would require completely changing the circuit itself.

## DEEP LEARNING AT THE PHYSICAL LAYER: SYSTEM REQUIREMENTS AND CHALLENGES

The target of this section is to discuss existing system-level challenges in PHY-DL, as well as the state of the art in addressing these issues. For a more detailed compendium of the state of the art, the interested reader can take a look at the following comprehensive surveys [6–9].

To ease the reader into the topic, we summarize at a very high level the main components and operations of a learning-based wireless device in Fig. 3. The core feature that distinguishes learning-based devices is that digital signal processing (DSP) decisions are driven by deep neural networks (DNNs). In particular, in the receiver

While DL can concurrently recognize the three parameters, traditional learning requires different feature extraction processes for each of the classification outputs. This, in turn, increases hardware consumption and hinders fine-tuning of the learning model.

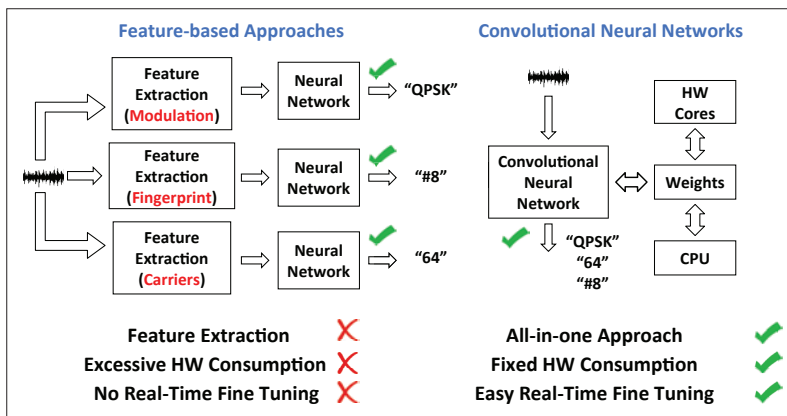


Figure 2. Feature-based approaches vs. convolutional neural networks.

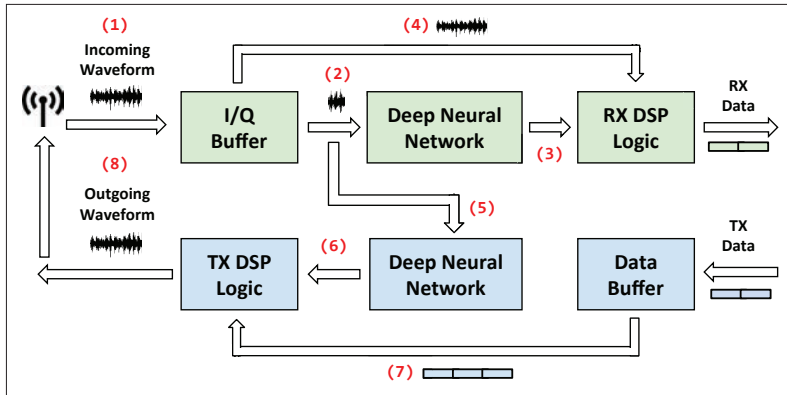


Figure 3. Main operations in a learning-based wireless device.

(RX) DSP chain, the incoming waveform is first received and placed in an I/Q buffer (step 1). Then a portion of the I/Q samples are forwarded to the RX DNN (step 2), which produces an inference that is used to reconfigure the RX DSP logic (step 3). For example, if a quadrature phase shift keying (QPSK) modulation is detected instead of binary PSK (BPSK), the RX demodulation strategy is reconfigured accordingly. Finally, the incoming waveform is released from the I/Q buffer and sent for demodulation (step 4). At the transmitter's side, the I/Q samples are sent to the RX DNN and to the transmitter (TX) DNN to infer the current spectrum state (e.g., spectrum-wide noise/interference levels). As soon as the inference is produced and the TX DSP logic is changed (step 6), the TX's buffered data is released (step 7), processed by the TX DSP logic, and sent to the wireless interface (step 8).

We identify three core challenges in PHY-DL, which are discussed below.

#### ADDRESSING LATENCY AND SPACE CONSTRAINTS

Domains such as computer vision usually do not have extreme requirements in terms of maximum latency or number of weights of a DL model. This is also true when ML is applied to higher layers of the protocol stack. For example, when we are uploading a picture on a social network, we do not expect a face recognition algorithm that automatically “tags” us and our friends to run under a given number of milliseconds. The same happens when running a routing protocol, where a few microseconds do not necessarily impact the protocol's performance. Very different, howev-

er, is the case of PHY-DL, where DSP constraints and hardware limitations have to be heeded — in some cases, down to the clock cycle level.

**System Challenges:** First, the DNN must run quickly enough to avoid overflowing the I/Q buffer and/or the data buffer (Fig. 3). For example, an incoming waveform sampled at 40 MHz (e.g., a WiFi channel) will generate a data stream of 160 MB/s, provided that each I/Q sample is stored in a 4-byte word. With an I/Q buffer of 1 kB, the DNN must run with a latency less than 6.25  $\mu$ s to avoid buffer overflow.

Moreover, the DNN must be fast enough to be (much) less than the channel's coherence time and the transmitter's frequency in changing parameters. For example, if the channel coherence time is 10 ms, the DNN should run with latency *much* less than 10 ms to make meaningful inference. However, if the transmitter switches modulation every 1 ms, the DNN has to run with latency less than 1 ms if it wants to detect modulation changes. The examples clearly show that lower DNN latency implies higher admissible sampling rate of the waveform, and thus, higher bandwidth of the incoming signal as well as higher capability of analyzing fast-varying channels.

Hardware resource utilization is a spiny issue. Nowadays, DL models usually have tens of millions of parameters; for example, AlexNet has some 60M weights and VGG-16 about 138M. Obviously, it is hardly feasible to entirely fit these models into the hardware fabric of even the most powerful embedded devices currently available. Moreover, it is not feasible to run them from the cloud and transfer the result to the platform due to the additional delay involved. Therefore, PHY-DL also has to be relatively small to be feasibly implemented on embedded devices. Resource utilization also directly impacts energy consumption, which is a critical resource in embedded systems. Indeed, the more area (i.e., lookup tables, block RAMs, etc.) the model occupies in the hardware, the higher the energy consumption. However, it has been shown [2] that thanks to the lower latency, implementing the model in a field-programmable gate array (FPGA) can lead to substantial energy reduction (up to 15 $\times$ ) with respect to an implementation in the central processing unit (CPU).

**Existing Work:** In [2], the authors propose *RFLearn*, a hardware/software framework to integrate a Python-level CNN into the DSP chain of a radio receiver. The framework is based on high-level synthesis (HLS) and translates the software-based CNN to an FPGA-ready circuit. Through HLS, the system constraints on accuracy, latency, and power consumption can be tuned based on the application. As a practical case study, the authors train several models to address the problem of modulation recognition, and show that latency and power consumption can be reduced by 17 $\times$  and 15 $\times$  with respect to a model running in the CPU. Moreover, it is shown that accuracy of over 90 percent can be achieved with a model of only about 30,000 parameters. Deep reinforcement learning (DRL) techniques are integrated at the transmitter's side with *DeepWiERL* [1], a hybrid software/hardware DRL framework to support the training and real-time execution of state-of-the-art DRL algorithms



on top of embedded devices. Moreover, DeepWiERL includes a novel supervised DRL model selection and bootstrap technique that leverages HLS and transfer learning to orchestrate a DNN architecture that decreases convergence time and satisfies application and hardware constraints.

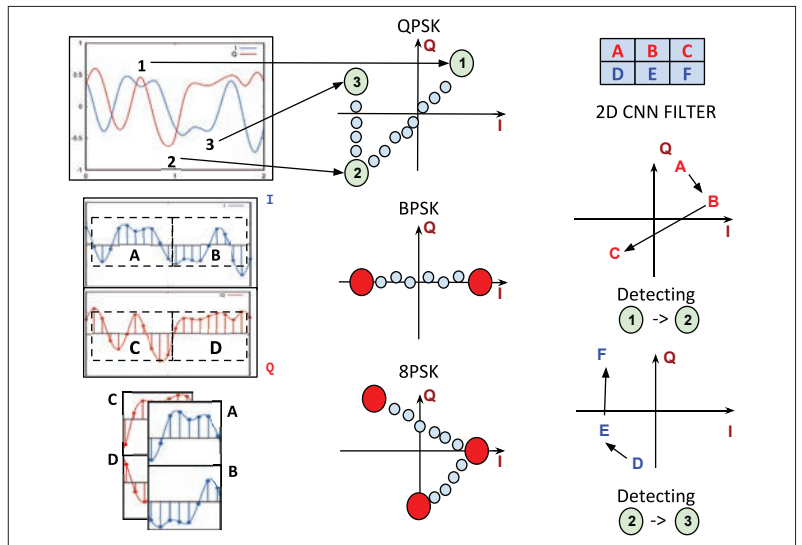
### DESIGNING FEATURES AND ADDRESSING STOCHASTICITY

In computer vision, DNNs are trained to detect small-scale “edges” (i.e., contours of eyes, lips, etc.), which become more and more complex as the network gets deeper (i.e., mouth, eyes, hair type, etc.). This is precisely the property that makes these networks excellent at detecting, for example, an object or a face in an image, irrespective of where it occurs. In the wireless domain, CNNs do not operate on images but on I/Q samples; thus, input tensors must be constructed out of I/Q samples.

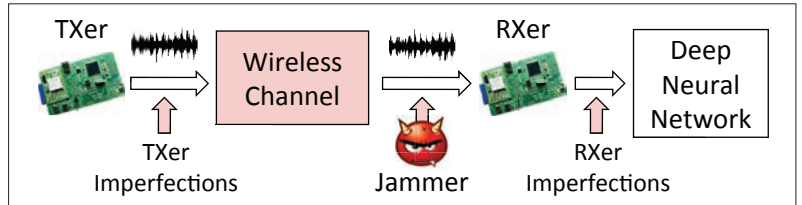
To make an example, the left side of Fig. 4 shows the approach based on two-dimensional (2D) convolution proposed in [2]. Specifically, input tensors are constructed by “stacking up”  $H$  rows of  $W$  consecutive I/Q samples. Figure 4 shows examples of transitions in the I/Q complex plane corresponding to QPSK, BPSK, and 8PSK. The transitions corresponding to points (1) to (3) are shown in the upper left side of Fig. 4. The figure clearly shows that different modulation waveforms present different I/Q transition patterns. For example, the transitions between (1, 0) and (-1, 0) peculiar to BPSK do not appear in QPSK, which presents a substantially different constellation. This can constitute a unique “signature” of the signal that can eventually be learned by the CNN filters. The right side of Fig. 4 shows an example of a  $2 \times 3$  filter in the first layer of a CNN trained for BPSK vs. QPSK modulation recognition. Specifically, the first row of the filter (i.e., A, B, C) detects I/Q patterns where the waveform transitions from the first to the third quadrant (which correspond to the symbol “1” to “2” transition in our example), while the second row (i.e., D, E, F) detects transitions from the third to the second quadrant (which correspond to the symbol “2” to “3” transition).

However, the above and similar CNN-based approaches [3] do not fully take into account that a PHY-DL system is inherently *stochastic* in nature. The first one is the unavoidable noise and fading that is inherent to any wireless transmission. Although channel statistics could be stationary in some cases, these statistics cannot be valid in every possible network situation; a CNN cannot be trained on all possible channel distributions and related realizations; and a CNN is hardly re-trainable in real time due to its sheer size. Recent research [4] has shown that the wireless channel makes it highly unlikely to deploy DL algorithms that will function without periodic fine-tuning of the weights [12]. Figure 5 summarizes the main sources of randomness in PHY-DL.

The second factor to consider is adversarial action (i.e., jamming), which may change the received signal significantly and usually in a totally unpredictable way. The third factor is the unavoidable imperfections hidden inside the RF circuitry of off-the-shelf radios (i.e., I/Q imbalance, frequency/sampling offsets, etc.). This implies that



**Figure 4.** On the left, we show how to construct an input tensor of size (in this example,  $10 \times 10 \times 2$ ) from an I/Q waveform. In the center, we show various examples of how a waveform corresponds to transitions in the I/Q complex plane for QPSK, BPSK, and 8PSK modulations. On the right side, we show how a  $2 \times 3$  filter of a CNN can distinguish the transition between the first, second, and third symbol of a modulation.



**Figure 5.** Source of randomness in PHY-DL.

signal features can (and probably will in most cases) change over time, in some cases in a very significant way.

**Existing Work:** The issue of PHY-DL stochasticity has been mostly investigated in the context of radio fingerprinting [13]. Specifically, the authors collected more than 7 TB of wireless data obtained from 20 bit-similar wireless devices over the course of 10 days in different environments. The authors show that the wireless channel decreases the accuracy from 85 percent to 9 percent. However, another key insight is that waveform equalization can increase the accuracy by up to 23 percent. To address the issue of stochasticity, the *DeepRadioID* system [4] was recently proposed, where finite input response filters (FIRs) are computed at the receiver’s side to compensate for current channel conditions by being applied at the transmitter’s side. The authors formulated a Waveform Optimization Problem (WOP) to find the optimum FIR for a given CNN. Since the FIR is tailored to the specific device’s hardware, it is shown that an adversary is not able to use a stolen FIR to imitate a legitimate device’s fingerprint. The *DeepRadioID* system was evaluated with a testbed of 20 bit-similar software-defined radios (SDRs), as well as on two datasets containing transmissions from 500 ADS-B devices and by 500 WiFi devices. Experimental results show that *DeepRadioID* improves the fingerprinting accuracy by 27 percent with respect to the state of the art.

Next-generation networks will necessarily require fast and fine-grained optimization of parameters at all the layers of the protocol stack. Radios will thus need to be extremely spectrum-agile, meaning that wireless protocols should be used interchangeably and according to the current spectrum circumstances.

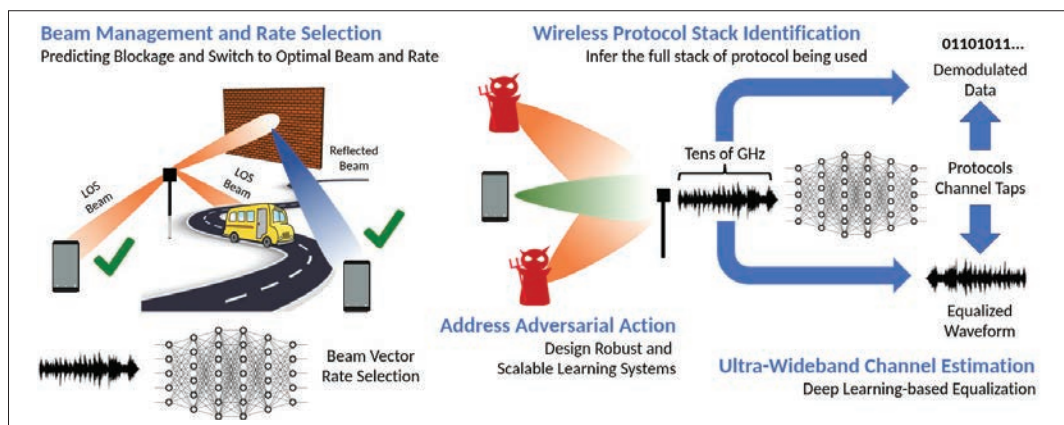


Figure 6. Summary of the main research challenges in PHY-DL and applications to 5GB networks.

## DEEP LEARNING AT THE PHYSICAL LAYER: THE WAY AHEAD

We now present an agenda of research opportunities in PHY-DL. Figure 6 summarizes the challenges discussed below.

### LARGE-SCALE EXPERIMENTS AND DATA COLLECTION

So far, PHY-DL techniques have been validated in controlled, lab-scale environments and with a limited number of wireless technologies. Although large-scale datasets in the area of radio fingerprinting have been produced, other PHY-DL problems (e.g., modulation recognition) have been clearly left behind. For this reason, the research community desperately needs large-scale experimentation to really understand whether these techniques can be applied in realistic wireless ecosystems where hundreds of nodes, protocols, and channels will necessarily coexist. Moreover, due to the current lack of common datasets, today every article in the wireless ML domain can claim to be “better than the previous one” in terms of accuracy. For this reason, the creation of large-scale datasets shared with the research community at large should also be considered as a priority.

To bring PHY-DL to the next level, we need “wireless data factories” able to generate I/Q data at an unseen scale. The newly developed Platforms for Advanced Wireless Research (PAWR) will be fundamental in addressing the above challenge (<https://advancedwireless.org>). The PAWR program will develop four platforms to be shared among the wireless research community. The platforms will enable sub-6 GHz, millimeter-wave, and drone experimentation capabilities in a multitude of real-world scenarios. Alongside PAWR, the Colosseum network emulator (<http://experiments.colosseum.net>) will soon be open to the research community and provide us with unprecedented data collection opportunities. Originally developed to support the Defense Advanced Research Project Agency’s (DARPA’s) spectrum collaboration challenge in 2019, Colosseum can emulate up to  $256 \times 256$  4-tap wireless channels among 128 software-defined radios. Users can create their own wireless scenarios and thus create “virtual worlds” where learning algorithms can be truly stressed to their full capacity.

## ADDRESSING WIRELESS ADVERSARIAL LEARNING

Up until now, researchers have focused on improving the accuracy of the PHY-DL model, without heeding security concerns. However, we know the accuracy of a DL model can be significantly compromised by crafting adversarial inputs. The first kind of attack is called *targeted*, where given a valid input, a classifier, and a target class, it is possible to find an input close to the valid one such that the classifier is “steered” toward the target class. More recently, researchers have demonstrated the existence of *universal perturbation vectors*, such that when applied to the majority of inputs, the classifier steers to a class different than the original one. On the other hand, the time-varying nature of the channel could compromise adversarial attempts. Moreover, the received waveforms still need to be decodable and thus cannot be extensively modified. Therefore, additional research is needed to fill the gap between adversarial machine learning (AML) and the wireless domain, and demonstrate *if*, *when*, and *how* AML is concretely possible in practical wireless scenarios.

### APPLICATIONS TO 5G AND BEYOND

Below, we discuss a series of applications of DL at the PHY layer to 5GB, and provide a roadmap of possible research avenues in the field.

**Analyzing Ultra-Wide Spectrum Bands:** The millimeter-wave (mmWave) and Terahertz (THz) spectrum bands have become the *de facto* candidates for 5GB communications. To fully unleash the power of these bands, mmWave/THz systems will operate with *ultra-wide spectrum bands* — on the order of several, perhaps tens of gigahertz (GHz). Thus, pilot-based channel estimation could not wind up being the best strategy. Indeed, frequently transmitting pilots for the whole bandwidth can lead to severe loss of throughput. A neural network could be trained to infer the channel directly based on the I/Q samples, without requiring additional pilots. One possible strategy could be to leverage the packet headers or trailers as source of reference I/Q data to train the learning model.

**Protocol Stack Identification:** Next-generation networks will necessarily require fast and fine-grained optimization of parameters at all the layers of the protocol stack. Radios will thus need to be extremely spectrum-agile, meaning that

wireless protocols should be used interchangeably and according to the current spectrum circumstances. To demodulate incoming waveforms transmitted with different strategies, it becomes necessary to infer the waveform type — and thus, the wireless protocol stack being used — before feeding it to the DSP logic. To the best of our knowledge, this problem remains open. Additional research should shed light on whether PHY-layer I/Q samples can be used to infer the whole stack of a wireless protocol. One possible avenue could be to extend the input size of the model and learn more complex features. However, this could increase latency to unacceptable levels. An alternative could be to utilize an ensemble model where smaller submodels are trained to analyze different portions of the waveform. This would ultimately help the model generalize while remaining under acceptable latency levels.

#### Blockage Prediction and Beam Alignment:

Another major challenge of mmWave and THz communications is the severe path and absorption loss (e.g., oxygen at 60 GHz). Moreover, mmWave and THz carriers cannot penetrate physical obstacles such as dust, rain, snow, and other opaque objects (people, building, transportation vehicles), making them highly susceptible to blockage. This key aspect will require high directionality of antenna radiations (i.e., beamforming), which will increase the transmission range but also introduce the compelling need for proactive beam steering and rate adaptation techniques. Deep learning could be utilized to design prediction techniques that can infer, in real time, an incoming blockage in a beam direction and thus proactively “steer” the beam toward another direction. In this spirit, Alrabeiah and Alkhateeb [14] have recently proven that under some conditions, we can leverage sub-6 GHz channels to predict the optimal mmWave beam and blockage status. Then the authors developed a DL model and tested it using a publicly available dataset called DeepMIMO. However, DeepMIMO is obtained through simulations based on a ray tracer, and sub-6 GHz channels may not always be available. Therefore, further research is needed to validate whether these approaches can be generalized to different channel conditions and obstacles.

Regardless of obstacles, transmitter (TX) and receiver (RX) beams have to be perfectly aligned to maximize the signal-to-noise ratio (SNR) during the transmission. Usually, through pilot sequences, the RX is then able to compute the SNR for each of the possible TX-RX beam combinations. The complexity of these beam alignment techniques is thus quadratic in the number of beams. A possible approach we are currently exploring is PHY-DL of ongoing transmissions between the TX and other receivers to infer the current TX’s beam and thus align the RX’s beam with the TX’s to avoid explicit beam scanning. We obtained some preliminary results through our mmWave testbed, where we train a CNN to identify the TX’s beam. We experimented with two 24-element phased array antennas, and with 12-beam and 24-beam codebooks. The results indicate that we are able to achieve accuracy close to 90 percent in both cases, with a CNN constituted by 7 convolutional layers (each with 64 kernels of size

$1 \times 7$ ) and 2 dense layers of 128 neurons, with a total of 848,472 parameters.

**PHY Virtualization and Optimization:** To deliver the required services, 5G will strongly depend on *virtualization* techniques, where PHY resources such as spectrum, transmission time, base stations, and so on will become shared among different virtual network operators (VNOs). This will allow seamless delivery of stringent quality of experience (QoE) requirements, such as real-time surveillance, web browsing, and high-quality video content delivery, among others. However, as the network size increases, the relationship between computing, storage and radio resources will be hard to model in explicit mathematical terms. To establish the resources, DRL could be utilized to learn representations of the current state of the system and tie them to optimal actions. Recently, the research community has started to move in this direction. Ayala-Romero *et al.* [15] presented a system where an autoencoder is used to project high-dimensional context data (e.g., traffic and signal quality patterns) into a lower-dimensional representation. Then an actor-critic neural network structure is used to map (encoded) contexts into resource control decisions. However, the proposed system is single-agent only, and the only PHY layer decision is related to the modulation and coding scheme used. It is unclear yet whether DRL can be extended to more complex problems and to multi-agents in realistic scenarios.

## CONCLUSIONS

The unprecedented scale and complexity of today’s wireless systems will necessarily require protocols and architectures to rely on data-driven techniques. In this article, we have provided an overview of PHY-DL and the state of the art on this topic. We have also introduced a roadmap of exciting research opportunities, which are definitely not easy to tackle but, if addressed, will take PHY-DL to the next level. We hope that this article will inspire and spur significant wireless research efforts in this exciting field in the years to come.

## REFERENCES

- [1] F. Restuccia and T. Melodia, “DeepWiERL: Bringing Deep Reinforcement Learning to the Internet of Self-Adaptive Things,” *Proc. IEEE INFOCOM*, 2020.
- [2] F. Restuccia and T. Melodia, “Big Data Goes Small: Real-Time Spectrum-Driven Embedded Wireless Networking through Deep Learning in the RF Loop,” *Proc. IEEE INFOCOM*, 2019.
- [3] T. J. O’Shea, T. Roy, and T. C. Clancy, “Over-the-Air Deep Learning Based Radio Signal Classification,” *IEEE J. Selected Topics in Signal Processing*, vol. 12, no. 1, Feb. 2018, pp. 168–79.
- [4] F. Restuccia *et al.*, “DeepRadiID: Real-Time Channel-Resilient Optimization of Deep Learning-based Radio Fingerprinting Algorithms,” *ACM MobiHoc*, 2019, pp. 51–60.
- [5] O. Naparstek and K. Cohen, “Deep Multi-User Reinforcement Learning for Distributed Dynamic Spectrum Access,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, 2019, pp. 310–23.
- [6] C. Zhang, P. Patras, and H. Haddadi, “Deep Learning in Mobile and Wireless Networking: A Survey,” *IEEE Commun. Surveys & Tutorials*, vol. 21, no. 3, 3rd qtr. 2019, pp. 2224–87.
- [7] H. Huang *et al.*, “Deep Learning for Physical-Layer 5G Wireless Techniques: Opportunities, Challenges, and Solutions,” *IEEE Wireless Commun.*, vol. 27, no. 1, Feb. 2019, pp. 214–22.
- [8] T. O’Shea and J. Hoydis, “An Introduction to Deep Learning for the Physical Layer,” *IEEE Trans. Cognitive Commun. and Networking*, vol. 3, no. 4, 2017, pp. 563–75.

The proposed system is single-agent only, and the only physical-layer decision is related to the modulation and coding scheme used. It is unclear yet whether DRL can be extended to more complex problems and to multi-agents in realistic scenarios.

- 
- [9] Q. Mao, F. Hu, and Q. Hao, "Deep Learning for Intelligent Wireless Networks: A Comprehensive Survey," *IEEE Commun. Surveys & Tutorials*, vol. 20, no. 4, 2018, pp. 2595–2621.
- [10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, 2015, p. 436.
- [11] T. D. Vo-Huu and G. Noubir, "Fingerprinting Wi-Fi Devices Using Software Defined Radios," *Proc. 9th ACM Conf. Security & Privacy in Wireless and Mobile Networks*, 2016, pp. 3–14.
- [12] G. Ditzler *et al.*, "Learning in Nonstationary Environments: A Survey," *IEEE Computational Intelligence Mag.*, vol. 10, no. 4, 2015, pp. 12–25.
- [13] A. Al-Shawabka *et al.*, "Exposing the Fingerprint: Dissecting the Impact of the Wireless Channel on Radio Fingerprinting," *Proc. IEEE INFOCOM*, 2020.
- [14] M. Alrabeiah and A. Alkhateeb, "Deep Learning for Mmwave Beam and Blockage Prediction Using Sub-6GHz Channels," *IEEE Trans. Commun.*, 2020.
- [15] J. A. Ayala-Romero *et al.*, "VrAI: A Deep Learning Approach Tailoring Computing and Radio Resources in Virtualized RANs," *Proc. ACM MobiCom*, 2019.

## BIOGRAPHIES

FRANCESCO RESTUCCIA [M'16] is an assistant professor at Northeastern University, Boston, Massachusetts. His research interests lie in wireless systems and artificial intelligence. He has published extensively in these areas and serves as TPC member and reviewer for top IEEE and ACM venues.

TOMMASO MELODIA [F'18] is the William Lincoln Smith Chair Professor at Northeastern University. He is the director of research for the PAWR Project Office. He is the Editor-in-Chief of *Computer Networks*, and an Associate Editor for *IEEE Transactions on Mobile Computing*, and *IEEE Transactions on Biological, Molecular, and Multi-Scale Communications*.