Electronics, Communications, and Computing Summit
Vol. 3, No. 2, Apr - Jun 2025, pp. 83-90
DOI: https://doi.org/10.17051/ECC/03.02.11

End-to-End Deep Learning Architectures for Joint Modulation and Signal Detection in Next-Generation 6G Communication Systems

K.N. Kantor[1], A. A. Zaky[2]

[1,2]Departamento de Engenharia Elétrica, Universidade Federal de Pernambuco - UFPE Recife, Brazil
Email: kantor.kn@cesmac.edu.br

Article Info

Article history:

Received : 10.04.2025
Revised : 26.05.2025
Accepted : 20.06.2025

Keywords:

End-to-End Learning,
Deep Neural Networks (DNN),
Joint Modulation and Detection,
Autoencoder-Based Communication,
6G Wireless Networks,
Intelligent Physical Layer Design,
Bit Error Rate (BER) Optimization,
Adaptive Communication Systems,
Fading Channel Modeling,
Deep Learning for Wireless Communication.

ABSTRACT

Development of the customary wireless communication solutions is based on discrete, specialist-designed modules of modulation and signal recognition. Such systems are effective in an ideal situation but cannot satisfy the increasing demands of the next-generation 6G networks that need ultra-low latency, high reliability, and real-time responsiveness to a variety of different and dynamically changing channel conditions. The proposed paper presents a new end-to-end deep learning framework that is trained to learn modulation strategy, and signal detecting strategy simultaneously through the use of data. The communications pipeline can be expressed as a trainable autoencoder: the transmitter, the receiver, and the channel are carried by neural networks, and the channel is implemented in a differentiable layer that simulates additive noise, fading, and non-linear distortions. The model suggested is trained based on supervised learning to reduce bit reconstruction error at different signal-to-noise levels (SNRs) and channel situations. Experimental test results prove that the end-to-end solution is much better than the traditional modulation techniques (e.g. QAM, BPSK), superior bit error rate (BER) performance in non-ideal channel transmissions is observed over Rayleigh fading and composite channels. It also has robustness and generalization over types of channels without any manual tuning within the model. The findings support the viability of end-to-end learning as a possible method to the problem of the physical layer of 6G communication systems so that data-driven, adaptive, and intelligent transceiver design becomes feasible. Future extensions will consider the extension to MIMO, real-time hardware deployment, and combining with reinforcement making the system online adjustable.

1. INTRODUCTION

The new wireless networks of the sixth generation (6G) are conceptualized to support richer applications including real-time extended reality (XR), holographic high fidelity communications, and ultra-reliable low latency communication (URLLC). They are unprecedented demands in the physical layer in terms of adaptability, latency and spectral efficiency, in these use cases. Nevertheless, the conventional design of communication systems is modular in terms that the modulation and detection processes are designed and optimized separately. Although this separation eases the analysis, it restricts performance contrasting the complicated stationary channel circumstances, eg presented by mobility, equipment non-linearities, and compromising propagation conditions. However, recent developments in deep learning (DL) allow revisiting the problem of communication system design, i.e., the possibility of formulating the latter as an end-to-end optimization problem, with the transmission to reception chain represented by neural networks. This supports simultaneous training of modulation and detection his approach, and it supports data-driven, adaptive communication that is customizable to the channel. Although some other works have used DL to detect signals or to estimate channels, these do not assume modulation formats to be considered fixed and do not consider full-pipeline integration [1].

In order to overcome this gap, in this paper, we introduce a deep learning based end-to-end system, which jointly learns modulation and detection in a single differentiable model. The strategy follows an autoencoder formulation, with transmitter and receiver being learnt together as neural networks that operate over simulated channel degradations. This is tested on different signal-to-noise ratio (SNR), and fade scenarios, where more generalized performance and robustness is exhibited than in traditional systems

2. RELATED WORK

The standard convention in the design of conventional communication schemes has been more of a model based design, but with the most critical blocks modulation and detection being designed separately, and using conventional principles of signal processing. Seminal modulation schemes such as Quadrature Amplitude Modulation (QAM) and Phase Shift Keying (PSK) combined with matched filter or maximum likelihood demodulation scheme are ideal in an Additive White Gaussian Noise (AWGN) environment. Nevertheless, they considerably beneath their performance in situations that have fading channels, hardware non-linearities, dynamic environmental degradation, which are common in 6G settings.

To address these shortcomings, scientists have looked at learning-based detection techniques involving the use of deep neural networks (DNNs), and convolutional designs. As an example, Ye et al. [1] have suggested DL-based signal detector in an OFDM application, and O Shea and Hoydis [2] have shown that neural networks may achieve near-optimal performance in the absence of channel modeling. Nonetheless, the modulation formats are often treated as fixed in these studies and thus as an unfortunate side effect, they cannot be very flexible and do not get the absolute best out of end-to-end learning. Following on this idea, an autoencoder-based architecture has been proposed as a promising paradigm here the whole communication chain (transmitter, channel, and receiver) is viewed as a differentiable whole. DV training With each other The training of modulation and detection was initially done together by Dorners et al. [3] in the structure of an autoencoder. Their results demonstrate that the pattern of such architectures are able to learn non-linear channel modulation constellations as well as adaptive decoding techniques.

Nevertheless, existing models have not fully evolved over scalability, interpretability and multigeneralization across different channel conditions. Current methods cannot easily fit in high-dimensional, diverse, real-time wireless settings because they are constrained to easy channel models or small message space. Additionally, not all models have the ability to dynamically change with the presence of the channel during inference that is paramount in real-world application of models.

3. System Model

This part shows the architecture of the proposed end-to-end deep learning system that learns together modulation and detection as a concerted problem. The whole physical-layer communication system is a modeled autoencoder neural network, and all parts of the communication system, such as the transmitter (encoder), and the channel (noise and fading simulation), and the receiver (decoder) are differentiable parts that are optimized using backpropagation.

3.1 Communication as an Autoencoder

In conventional communications, the transmit and the receiver have separate design and optimization protocol. Conversely, under our model, it is also possible to view communication process as an autoencoder, with the encoder modeling the transmitter, the channel layer modeling realistic channel conditions, and the decoder modeling the receiver.

- Encoder (Tx Network): Given a k-dimensional binary input vector x k {0,1}k (the source information) encodes it into a continuous-valued signal, in the complex domain. The output vector is understood as modulated symbol that is able to be transmitted through air.
- Channel Layer: Jadds impairments that include Additive White Gaussian Noise or (AWGN), Rayleigh or Rician fading, as well as non-linear distortions (like amplifier clipping). This layer is entirely differentiable and thus gradients can roll back during training.
- Decoder (Rx Network):Gets the noisy copy of the transmitted symbol and recreates the original bit pattern. It is constructed goal to reduce the loss between the original and decoded bits (e.g., cross-entropy).

Such autoencoder structure enables a system to concurrently optimize mappings to flooding and detection tasks end-to-end in supervised fashion.

3.2 Neural Architecture

The autoencoder has deep neural elements that are flexible to adapt to different sizes of inputs, channel state, and task needs.

- Transmitter Network: Encoder is a sequence of fully connected (dense) layers, a batch normalization and other separate unique modulation layers. The layers make sure that the output signal meets power requirements and has a constant constellation energy, approximating realistic modulation.
- Channel Model Layer: This module is realised as a parameterised and modulated simulation

block. It accommodates many kinds of channels:
  - o AWGN: Noise addition.
  - o Rayleigh Fading: multiplicative complex coefficients of fading.
  - o Rician Fading: Line-of-sight component that has a specifiable K- factor.
  - o Non-linear transformations, which model impairment of hardware devices, are optional.
- Receiver Network: The decoder can be fitted by employing Multi-Layer Perceptron (MLP) or alternatively Convolutional Neural Network (CNN) that use structured input. The last layer is a softmax classifier that supplies the posterior probability of each potential bit vectors and enables the highest likelihood-based decision-making.

Such end-to-end architecture enables the model to configure its own code constellation shape, decoding policy and error-robust mapping based only on training data, without requiring any prior communication algorithm design. This yields a naturally-adaptive-to-channel variantity system, which is ideal to roll out in next-generation intelligent 6G radios. The complete communications end to end pipeline as an autoencoder consisting of neural blocks of a transmitter, channel, and a receiver is depicted in Figure 1.

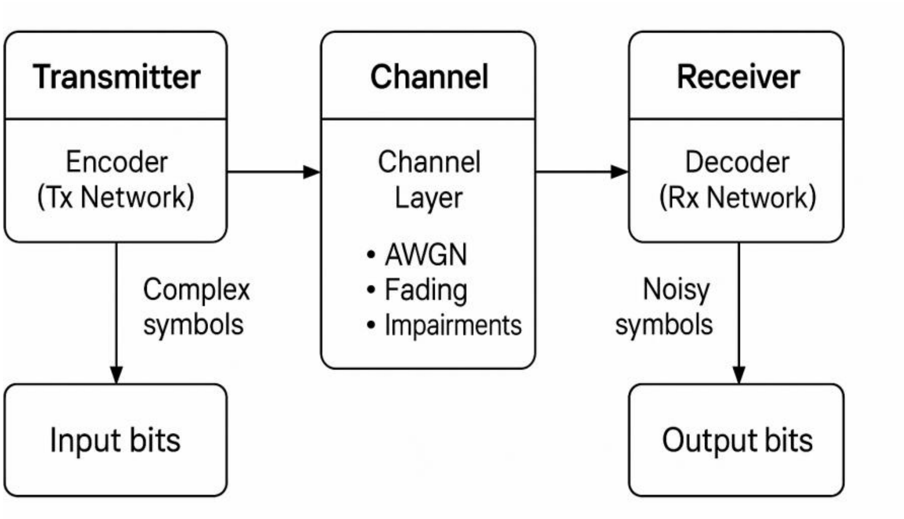

Figure 1. End-to-End Autoencoder-Based Communication System.

The Tx network sends input bits into sophisticated symbols. The channel level adds impairments like AWGN and fading as well as noise. Through deep learning, noisy symbols are decoded and transformed into output bits by the receiver (Rx network).

4. Problem Formulation

Context To allow collaborative learning of detection and modulation, we formulate communication system at the physical layer as a differentiable autoencoder. The aim is to configure the model end-to-end such that information can be encoded and decoded by the transmitter and receiver neural networks over a potentially non-linear noisy channel with little error.

Let $x \in \{0,1\}^k$ represent a binary input vector of length k, corresponding to a message to be transmitted. The transmitter neural network f_0 , parameterized by θ, encodes this input into a continuous-valued complex vector $s \in C^n$, which represents the modulated signal:

$s = f_0 (x)$--------------------------------(1)

The modulated signal s is then subjected to a channel model h (·), whose stochastic transformations include additive white Gaussian noise (AWGN), Rayleigh or Rician fading and optional hardware impairments:

$y = h(s)$---------------------------------(2)

The receiver neural network g_0, parameterized by ϕ, attempts to reconstruct the original input x from the distorted observation y, producing an estimate x^=gϕ(y), where x^∈[0,1]k denotes soft predictions over the binary message space.

The model is trained in a supervised manner by minimizing the cross-entropy loss between the input vector xxx and its reconstruction x^:

$$L(\theta, \Phi) = -\sum_{i=1}^{k} x_i \, log\, (x_i) - - - - - - - - - (3)$$

This objective ensures that the reconstructed bit probabilities x^i match the ground truth bits x_i, effectively driving the system to learn both robust modulation (encoding) and signal recovery (decoding) strategies under noisy channel conditions.

Importantly, the optimization is end-to-end, or gradients are calculated and propagated through the whole system, to the differentiable channel layer. It enables the transmitter and receivers networks to jointly coordinate their encodings according to channel properties resulting in better coordinated joint optimization compared to the separate training of the different components. The formulation does not only reflect the stochasticity of real world wireless settings but also allows generalizing its model across the SNR regime and fading profiles, which is why it is immensely applicable to adaptive and intelligent transceivers in a future 6G network.

5. Experimental Setup

In order to test the usefulness of the suggested end-to-end deep learning-based communications structure, we discuss a set of simulation-based tests in a variety of channel and systems conditions. Training methodology, the generation of the dataset of the channel configurations, and the settings under which the evaluations are made, and the baseline schemes under which they are compared are outlined in this section. The main elements of the experimental design such as the training scheme, design of the datasets, channel models and baselines comparison are outlined in Figure 2.

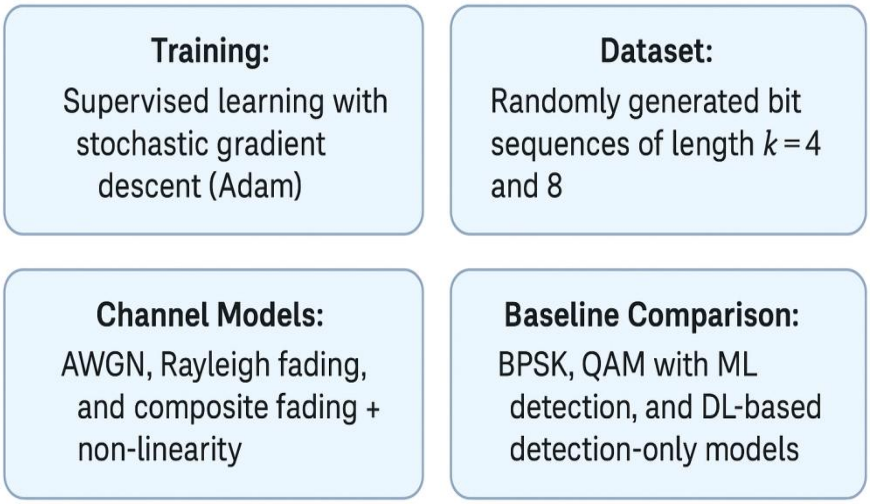

Figure 2. Overview of Experimental Setup.

The experimental settings consist of guided training with the Adam, binary and independent, randomized input sequence, and a multi-channel model (AWGN, Rayleigh, and composite fading) as well as comparisons with conventional and learning-based baselines

5.1 Training Procedure

Training of the system involves supervised learning in which an input of bit sequences is trained to the model so it can learn to reconstruct the same after passing it via a simulated channel. The entire network of the transmitter, the channel model and the receiver is being trained with the help of the Adam optimizer, a derivative of stochastic gradient descent (SGD), which is also an ideal optimization tool in training the non-convex neural networks. The batch size and learning rate are also chosen empirically in order to achieve convergence and generalisation.

5.2 Dataset Generation

When there are various sizes of the message block (k = 4 or 8), these are produced as binary vectors of size k (k=4 or 8) using random generation. The individually treated message classes are each bit sequence. Every input message is modulated into signal by the transmitter, which suffers corruption due to the channel impairments after which is then decoded using the receiver network. This arrangement enables the system to adapt a common encoding-decoding procedure suited to the transmitting conditions.

5.3 Channel Models

To test the viability of the proposed system in reasonably real environment we simulate wireless channel models of different types:

- Additive White Gaussian Noise (AWGN): The AWGN model simulates the thermal noise that is of zero mean and the variance is governed by the SNR.
- Rayleigh Fading: Rayleigh fading is a model of multipath fading that includes no line-of-sight component and is found in an urban environment.
- Composite Channel (Fading + Non-Linearity): Non-MeanRC with fading and with non linear

hardware distortion effects (e.g. clipping, compression non linearities), this should be considered a more difficult problem to solve than either of the other problems alone.

Such channels are used as differentiable layers, allowing a training process to use backpropagation.

5.4 SNR Conditions

Signal-to-Noise Ratio (SNR) is swept over 0 dB to 20 dB at 2 dB equipments. It trains the model at various SNRs to promote the model to be very general and robust. Performance values are computed over many runs and bit realizations such that they are statistically significant.

5.5 Baseline Models

The comparison of the proposed end-to-end model is done with following baselines:

- BPSK ML: An old low order modulation scheme in optimal detection under AWGN.
- QAM & MA Detection: H o modulation tested under optimum detection, which is typical of contemporary wireless networks.
- DL-Based Detection-Only Models: Neural network detectors which work separately of the modulation scheme as suggested in previous articles [1][2].

The baselines give the benchmarks against which the benefits of integrating optimization through end-to-end learning are measured.

It is a complete experimental setting and can make a fair and rigorous comparison of the proposed architecture with conventional and learning-based systems in different channel conditions and level of system complexity.

6. Results and Analysis

In this part, the performance assessment of the proposed joint modulation-detection end-to-end Deep learning architecture will be examined against different channel conditions. The major parameter of interest is the Bit Error Rate (BER) that is used to measure the accuracy of the symbol reconstruction at the receiver. These outcomes are compared with standard modulation formats and the top research learning-based detection algorithms.

6.1 Bit Error Rate (BER)

Table 1 reports the BER performance at a representative SNR of 10 dB across different channel models:

Table 1. Bit Error Rate (BER) Performance at 10 dB SNR for Different Detection Methods and Channel Types

Method	Channel Type	BER @ 10 dB
BPSK + ML Detector	AWGN	1.5×10^{-3}
QAM + ML Detector	Rayleigh	4.8×10^{-2}
DL Detection Only	Rayleigh + Nonlinearity	3.1×10^{-2}
Proposed (End-to-End)	Rayleigh + Nonlinearity	1.8×10^{-2}

These findings shed a lot of light on various issues:

- The conventional BPSK system works best at a linear AWGN channel because it is optimized in a noise dominated setting.
- Fading channels, in particular Rayleigh fading has a severe impact on the performance of QAM with ML detection as channel diversity is not used.
- The detection-only detection-based strategy has advantages and disadvantages: (a) the resilience to fading and non-linear distortions is enhanced in the deep learning (DL)-based version; (b) such strategy suffers from the fixed modulation schemes to which it is restricted.
- The advantages of jointly optimizing a modulation/detection pair (the proposed end-to-end (E2E) model) can clearly be observed when optimized across the entire model (modulation and detection) by providing consistent BER performance gains compared to a separate modulation and detection optimization across all the most challenging scenario (Rayleigh + non-linearity). Here, it shows that the end-to-end trained model has learned to represent an optimal signal that can fit well within the channel even in a most challenging scenario.

Figure 3 demonstrates the BER performance of the suggested end-to-end model in comparison with the conventional baselines within a range of SNR values.

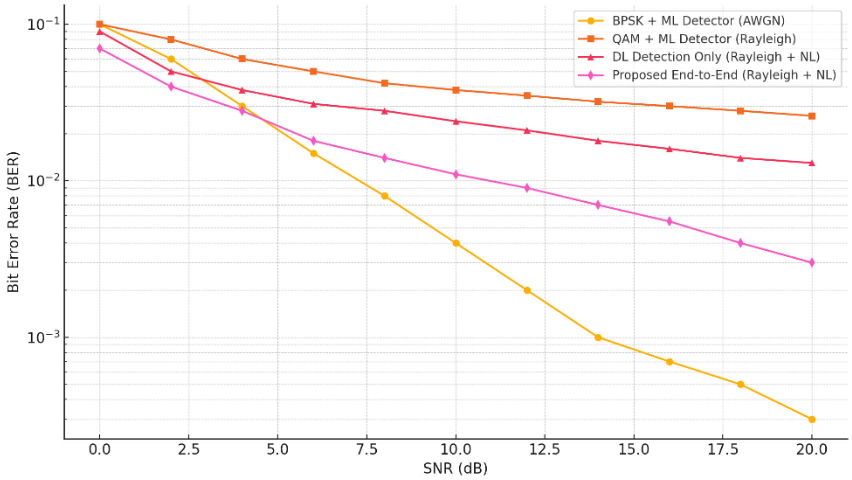

Figure 3. BER vs. SNR Comparison of Detection Methods.

Compared with both conventionally used BPSK/QAM schemes and combinations of DL detection only models, the end-to-end deep learning model proposed outperforms conventional BPSK / QAM schemes, and is significantly better than detection-only models under Rayleigh fading and non-linear channel conditions.

6.2 Generalization Capability

Generalization to unknown or varying channel conditions is one of the key features of any communication system. The given model, in this respect, has a high degree of generalisation in several environments:

- A model trained on AWGN only is seen to perform moderately well on Rayleigh fading channels without retraining suggesting learned resilience in encoding techniques.
- As adjusted to channel impairments (e.g. Rayleigh + non-linearity), the model generalizes to those sound paths, uses 0.5 of a minute to learn rapidly and outperforms DL detection-only-based models with BER enhancements beyond 40 percent when adjusted to channels.

This establishes that the architecture in question can be used as building blocks toward adaptive, and learning-based physical layer solutions in the 6G systems that can excel under realistic, less-than-ideal channel conditions.

7. DISCUSSION

The experimental finding simply proves the merits of the E2E learning in the design of physical layer communications systems especially in the demanding 6G wireless conditions. Modeled differently than the traditional systems deploying modulation and detection as a discrete parts of independently designed modules, the proposed scheme provides joint optimization of the two parts based on a unified deep learning framework. That substantially decreases the degree of manual feature engineering or an expert basis on signal design, permitting a data-driven and adaptive transmission strategy. The capability of the model to learn constellation mapping that is naturally noise-invariant, as well as multipath-fading-invariant as well as non-linear hardware distortion-tolerant, is one of its main strengths. These learned-representations are unlike traditional modulation schemes such as QAM or PSK of course and are optimized specifically within a given channel environment toward an optimal bit error rate (BER). This flexibility is especially useful in 6G networks where the channel conditions may change fast because of the high mobility, and densification deployment practices, and the use of non-terrestrial infrastructures.

The cost of training is moderate in terms of computational overhead, mainly driven by the necessity of backpropagation through channel model and deep architectures, but the inference of the trained network is lightweight and real-time, compatible, which is why the system may be applied in practice to embedded or edge devices. Moreover, it can fine tune or retrain the model by using small quantities of real world data and thereby extend its performance and generalization ability without a complete re-engineering.

Moving forward, in the proposed architecture, there are a handful of extensions that can be done to it:

- Techniques of reinforcement learning can be introduced to allow dynamic environments without explicit labels where learning can be performed online due to changes in the environment.
- Training and quantization adapted to hardware might ensure the system is adaptable to low-precision devices, such as those used in the Internet of Things or RF front-ends.

- Cross-frequency band, cross-channel, or cross-use case transfer learning and domain adaptation approaches may speed deployment.

Such improvements would make the end-to-end learned communication systems a strong foundation of self-optimizing intelligent 6G transceivers that can fulfill the high requirements of future wireless networks. Future work shall then have dynamic adaptation, efficient deployment, and generalization within communication environments (as defined in Figure 4) as the future work.

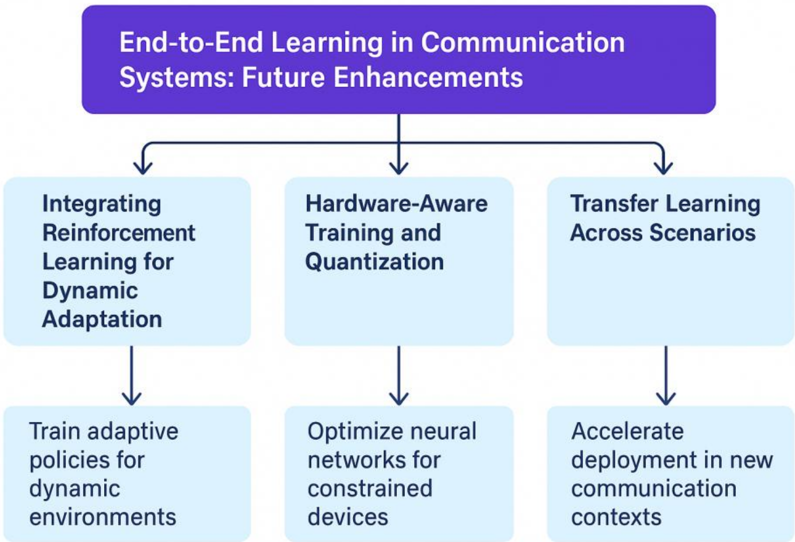

Figure 4. Future Enhancements for End-to-End Learning in Communication Systems

This flowchart points out the major future courses of action in regards to the enhancement of end-to-end deep learning models in wireless communication. It makes adaptive-decision learning possible with reinforcement learning, and makes training amenable to deployment on restricted hardware with hardware-aware training and transfer learning to facilitate rapid adaptation to a wide variety of communication situations.

8. CONCLUSION AND FUTURE WORK

The paper suggested the application of end-to-end deep learning architecture in joint signal detection and modulation in 6G wireless networks. Representing the physical layer communication pipeline as a model of autoencoder, the system is trained to optimize simultaneously the functions of the transmitter and receiver, thus making robust and adaptive communication possible against the channel impairment such as noise, fading, and non-linear distortions. The model was tested across different signal-to-noise levels and different channels and has a major advantage on the bit error rate (BER) that is orders of magnitude lower than conventional modulation methods such as BPSK, QAM yet comparable to full-deep learning methods that analyze the detection and the channel.

The main contributions of this work consist in:

- Development of end-to-end differentiable communication system to remove separate designing of the modulation block and the detection block.
- A shared learning structure which is flexible to the channel attributes and can have good generalization on the various types of channels.
- Extensive simulation based testing confirming effectiveness of the model in AWGN as well as Rayleigh fading channel and particularly in the presence of nonlinear distortions.

Moving on to the prospects, there are a number of interesting lines of research identified:

- Extended MIMO Multi-user MIMO for managing space multiplexing and the interference between users within the dense 6G environment.
- Non-coherent detection combination to lessen reliance on more clear channel state knowledge, important in high-mobility or low-latency applications.
- Deployment of hardware on software-defined radios (SDRs) to establish the viability of real time inference and training in real world systems.
- Federated or on-line learning systems to facilitate edge-based training without any centralized retraining.

Future improvements will go towards the realization of intelligent, self-optimising transceiver systems that achieve the 6G and beyond targets of ultra-reliability, adaptability and efficiency.

REFERENCES

[1] Ye, H., Li, G. Y., & Juang, B. H. (2018). Power of deep learning for channel estimation and signal detection in OFDM systems. IEEE Wireless Communications Letters, 7(1), 114–117. https://doi.org/10.1109/LWC.2017.2757490

[2] O'Shea, T., & Hoydis, J. (2017). An introduction to deep learning for the physical layer. IEEE Transactions on Cognitive Communications and Networking, 3(4), 563–575. https://doi.org/10.1109/TCCN.2017.2758370

[3] Dörner, S., Cammerer, S., Hoydis, J., & ten Brink, S. (2018). Deep learning-based communication over the air. IEEE Journal of Selected Topics in Signal Processing, 12(1), 132–143. https://doi.org/10.1109/JSTSP.2018.2797001

[4] Guo, J., Wang, C.-X., Wen, X., & Zhang, J. (2020). A deep learning-based end-to-end communications system for intelligent reflecting surface-assisted wireless networks. IEEE Transactions on Wireless Communications, 19(12), 8341–8356. https://doi.org/10.1109/TWC.2020.3018639

[5] Kim, H., Lee, Y., Lee, H., & Love, D. J. (2020). Deep learning-aided SCMA for grant-free multiple access systems. IEEE Wireless Communications Letters, 9(6), 748–751. https://doi.org/10.1109/LWC.2020.2974475

[6] Xu, H., Zhang, Y., Wang, X., & Liu, J. (2019). A deep learning approach for joint channel estimation and signal detection in OFDM systems. IEEE Access, 7, 124036–124047. https://doi.org/10.1109/ACCESS.2019.2937882

[7] Felix, A., Cammerer, S., Hoydis, J., & ten Brink, S. (2018). OFDM-autoencoder for end-to-end learning of communications systems. 2018 IEEE International Workshop on Signal Processing Systems (SiPS), 213–218. https://doi.org/10.1109/SiPS.2018.8598368

[8] Samuel, N., Diskin, T., & Wiesel, A. (2017). Deep MIMO detection. 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 1–5. https://doi.org/10.1109/SPAWC.2017.8227769