Implicit Channel Learning for Machine Learning Applications in 6G Wireless Networks

Ahmet M. Elbir, Wei Shi, Kumar Vijay Mishra, Anastasios K. Papazafeiropoulos, and Symeon Chatzinotas

***Abstract*—With the deployment of the fifth generation (5G) wireless systems gathering momentum across the world, possible technologies for 6G are under active research discussions. In particular, the role of machine learning (ML) in 6G is expected to enhance and aid emerging applications such as virtual and augmented reality, vehicular autonomy, and computer vision. This will result in large segments of wireless data traffic comprising image, video and speech. The ML algorithms process these for classification/recognition/estimation through the learning models located on cloud servers. This requires wireless transmission of data from edge devices to the cloud server. Channel estimation, handled separately from recognition step, is critical for accurate learning performance. Toward combining the learning for both channel and the ML data, we introduce *implicit channel learning* to perform the ML tasks without estimating the wireless channel. Here, the ML models are trained with channel-corrupted datasets in place of nominal data. Without channel estimation, the proposed approach exhibits approximately 60% improvement in image and speech classification tasks for diverse scenarios such as millimeter wave and IEEE 802.11p vehicular channels.**

***Index Terms*—Machine learning, channel estimation, artificial intelligence, wireless communications.**

I. Introduction

Lately, the fifth generation (5G) wireless networks are very close to operational deployment. The lessons learned from the previous and ongoing 5G research is paving way for conceptualization of 6G technologies within the wireless communications community [1]. Different from the previous generation, the 6G networks are envisioned to enhance and aid emerging applications such as virtual and augmented reality (VAR), autonomous vehicles (AVs) and artificial intelligence (AI) together with ubiquitous connectivity and massive number of devices. In particular, 6G is expected to achieve higher peak data rates ($>$ 100Gb/s), lower latency ($<$ 1ms) and enhanced reliability (99.999%) [1, 2]. Furthermore, compared to 5G systems incorporating massive multiple-input multiple-output (MIMO) configurations, 6G networks will operate at higher frequencies (up to 10 THz) to exploit high-capacity data transmission with wider bandwidth and ultra-massive (UM) MIMO architectures.

In order to design the physical layer in 5G networks, recently machine learning (ML) techniques have been introduced for the challenging design problems, e.g., resource management [3, 4], symbol detection [5], beamforming [6] and channel estimation [7]. While these methods have shown a great potential for system design to deal with data, hardware, and computational complexities, they have not been deployed in the current 5G architectures. Rather, the ML-based techniques are envisioned as the primary candidate to be considered in 6G network design [1, 2, 8].

In the upcoming 6G era, the wireless network will incorporate massive number of devices such as mobile phones, connected vehicles, drones, and internet of thing (IoT) devices. Hence, there will be a tremendous increase in the amount of data generated by these devices. The International telecommunications union (ITU) estimates that the global mobile traffic in 2030 will reach 5016 exabytes (EB), of which more than 70% will be image or video [9]. Moreover, AVs are expected to generate approximately 20 TB/day/vehicle data, a part of which is transmitted to the road side infrastructure [10]. To process and extract useful information from these images/videos, several ML techniques have emerged as a key enabling technology in the field of image classification, face/object/motion detection, target tracking, and speech recognition (Fig. 1). These applications require huge amount of data to be processed and learned by an ML model for extracting the features from the raw data and provide a "meaning". As a result, most of the processed data in 6G is expected to be generated or processed by ML algorithms.

The performance of the ML models depends on the data collected from the edge devices in the network. However the wireless channel corrupts ML data (image/video/speech) during transmission and reduces the inference performance [5]. Thus, wireless channel acquisition is a crucial task in the ML applications. Furthermore, the 6G requirements, e.g., ultra-low latency and high mobility, make the channel acquisition process even more challenging [1]. To this end, several model-based [11, 12] and ML-based [6, 7, 13] channel state information (CSI) estimation techniques have been proposed for various communications standards in cellular (5G millimeter wave (mmWave)) and vehicular (IEEE 802.11p) networks. The model-based techniques have high energy and time consumption because they usually require tackling a high-dimensional optimization problem, especially for UM-MIMO configuration. On the other hand, the ML-based methods are

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), Ericsson Canada and the ERC Project AGNOSTIC.

A. M. Elbir is with the Department of Electrical and Electronics Engineering, Duzce University, Duzce, Turkey, and the SnT at the University of Luxembourg, Luxembourg (e-mail: ahmetmelbir@gmail.com).

W. Shi is with the School of Information Technology, Carleton University, Ottawa, Canada (e-mail: wei.shi@carleton.ca).

K. V. Mishra is with the United States Army Research Laboratory, Adelphi, MD 20783 USA (e-mail: kvm@ieee.org).

A. K. Papazafeiropoulos is with the CIS Research Group, University of Hertfordshire, Hatfield, U. K., and SnT at the University of Luxembourg, Luxembourg (e-mail: tapapazaf@gmail.com).

S. Chatzinotas is with the SnT at the University of Luxembourg, Luxembourg (e-mail: symeon.chatzinotas@uni.lu).

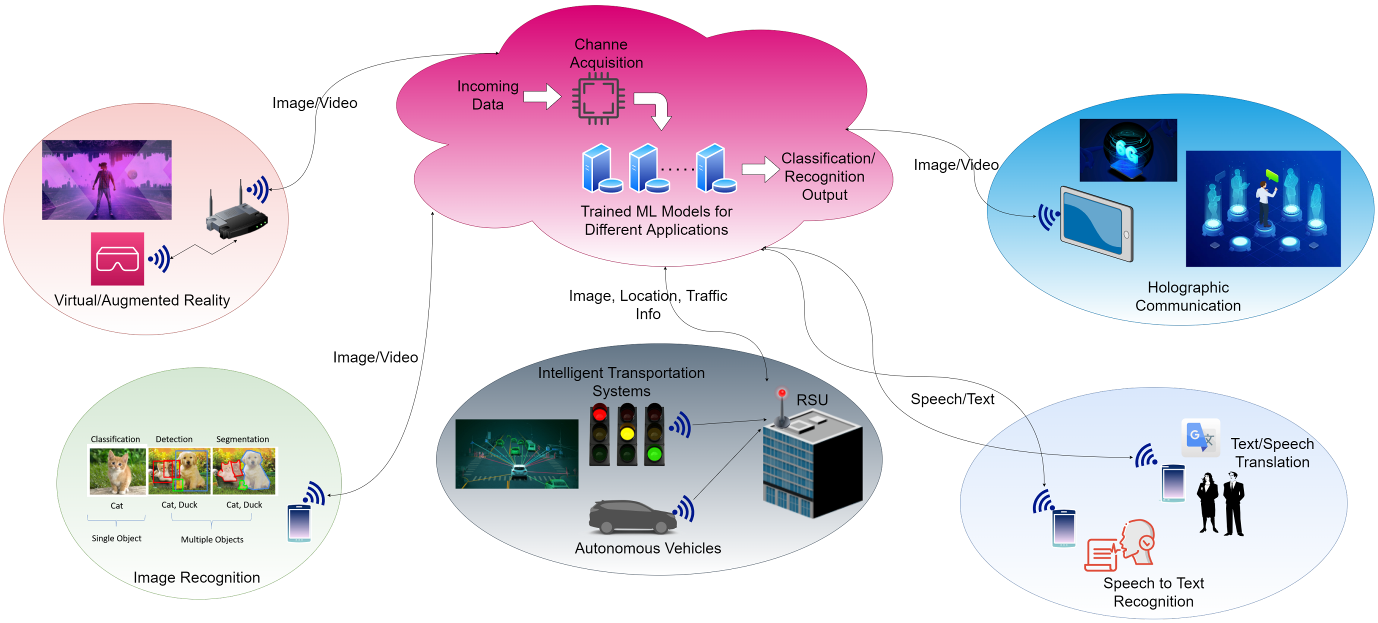

Fig. 1. Various ML applications for 6G use-cases and the corresponding transmit data types. The edge devices send the ML data to the PS through wireless medium. Once the channel acquisition is performed, the received ML data is input to the learning models. The recognition task is carried out at the PS that sends back the prediction output to the edge devices.

TABLE I
STATE-OF-THE-ART FOR CHANNEL ESTIMATION TECHNIQUES

Channel estimation	**Algorithm**	**Advantages**	**Drawbacks**
Model-based for massive MIMO [11]	Codebook-based adaptive compressed sensing	Simple implementation especially for single-path channels	Inferior performance because of grid-based channel construction
Model-based for IEEE 802.11p [12]	LS estimation with constructed data pilots	Improved representation of the data correlation characteristics	High computational complexity
ML-based for IEEE 802.11p [13]	DNN design with spectral-temporal averaging	Low computational complexity	Performance is data-dependent and worsens for new data
ML-based in massive MIMO [6]	Offline and online deployment with hybrid beamforming	Implicit via CNN fed with received pilots	Limited performance at high SNR because of low resolution
ML-based in massive MIMO [7]	CoNN design with spatial-frequency-temporal correlation	Low complexity for wideband systems	Requires full matrix inversion for input design
Implicit for massive MIMO [4]	MLP model with codebook-based beam training	Bypassing the channel estimation stage for hybrid beamforming	Requires prior, e.g., user locations and environment geometry
Implicit for massive MIMO [5]	Conditional GAN model	Bypassing the channel estimation stage for signal detection	Requires prior, e.g., received pilot signals
Implicit for IRS-assisted massive MIMO [14]	Graph neural network architecture for beamformer design	Bypassing the channel estimation stage to obtain IRS beamformers	Requires prior, e.g., user locations
Implicit for ISAC [15]	Path loss and Doppler shift estimation via grid search	Bypassing the channel estimation stage for transmitter design	Requires prior, e.g., location and speed of the vehicles

data-driven, learn the features of the raw data, and provide robustness against noise and channel corruptions. Thus, it is instructive to leverage ML to address the uncertainty in channel estimates. However, implementation of ML techniques necessitates the task-oriented design of multiple ML models, each of which is dedicated to a different layer in the Open System Interconnection (OSI) communications model [2]. For instance, separate ML models for physical (e.g., channel estimation) and application (e.g., image recognition) layer tasks should be devised.

Motivated by the aforementioned two 6G facets — ML-related data and ML-based network design — we introduce *implicit channel learning* approach by combining the learning for both wireless channel and ML-related data. To this end, an ML model is trained with the dataset carrying the accompanied wireless channel such as path loss and correlation. Thus, the trained model implicitly learns the channel characteristics and makes accurate predictions even if the ML data are corrupted by the wireless channel. While there are a few studies on the implicit estimation of wireless channel for symbol detection [5] and active/passive beamforming [4, 14, 15], these methods require either *a priori* information such as user locations, or received pilot signals along with the optimization of the remaining system parameters, e.g., beamformers (see Table I). Unlike these prior works, our proposed approach addresses both channel learning and application layer learning tasks through a single ML model, which is fed with the channel corrupted image/speech data. Thus, it is helpful for future 6G networks that are likely to heavily utilize ML-related data.

Our work is connected with the rich body of signal processing literature on classic problems of *blind channel estimation*, *blind equalization*, and *blind decoding*, where the channel is unknown and not estimated *a priori* while decoding the communications messages. Instead, our technique predicts the classification/recognition information of the transmitted data.

In the next section, we discuss common applications of ML in wireless communications. Next, we survey the state-of-the-art in channel estimation for various network architectures such as cellular and vehicular networks (see Table I). In this context, we introduce our implicit channel learning approach and validate its performance on image and speech classification tasks under the effect of various wireless standards. Finally, we identify the major design challenges in realizing implicit channel learning and highlight some future research directions.

II. ML Applications in Wireless Networks

In many contemporary ML applications, training and prediction of the ML model are carried out at a central parameter server (PS) of the network, whereas the data are generated at the network edge comprising devices like mobile phones, vehicles, and IoT sensors. These ML models are composed of huge number of parameters. For instance, well-known ML models such as AlexNet, VGG (both for image classification), and GPT-3 (text recognition/translation) are comprised of 60 million (240 MB), 136 million (552 MB), and 170 billion (680 GB) learnable parameters, respectively [8, 9]. Storing these huge models in the edge devices is costly and inefficient. Instead, these models are stored at cloud PSs. The training stage is managed offline using a pre-collected training dataset. But the prediction-stage classification/recognition demands transmission of the generated image/video/speech to the PS because the ML models run short of storage at the edge devices. Once the prediction is performed at the PS, its result is sent back to the edge device (Fig. 1). For example, Google's image recognition technology Lens transmits captured images via a smart phone to the PS, where a pre-trained convolutional neural network (CNN) used for recognition/classification/translation brings up the information related to the objects in the captured image. Similarly, the AVs have on-board pre-trained learning models but still perform data transmission (reception) to (from) the cloud platforms for map generation, path planning, and forecasting [10].

Fig. 2 illustrates the impact of channel estimation on ML applications of image and speech classification for mmWave [11] and vehicle-to-vehicle (V2V) [12] line-of-sight (LoS) links. We used MNIST and Speech Command datasets for image and speech classification tasks, respectively, each of which has 10 classes [9]. The image dataset includes the black-and-white handwritten digits whereas the speech dataset is composed of the spectrogram of audio signals. During training, *clean* (with equalized channel) datasets are used for both ML models while

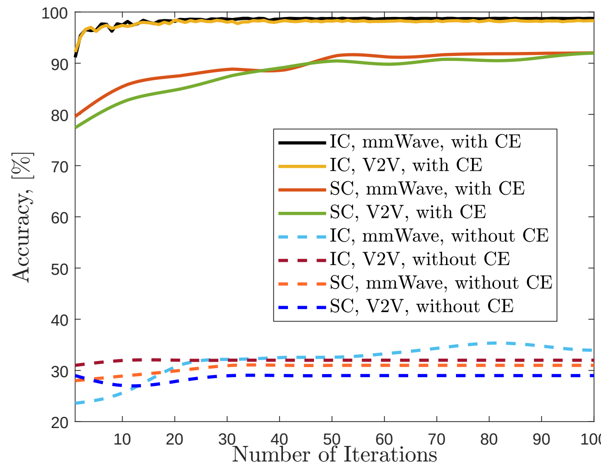

Fig. 2. Impact of channel estimation (CE) for image (IC) and speech classification (SC) performance on the validation datasets. The details for the datasets and the learning models are given in Table II.

two different validation datasets are prepared with (*clean*) and without (*corrupted*) channel estimation to observe its effect on learning accuracy. We observe that the classification performance degrades significantly if the images/spectrograms are corrupted by the wireless channel without equalization for both tasks. This shows the significance of the channel acquisition in ML applications.

III. Channel Estimation Techniques

The model-based signal processing techniques need accurate mathematical modeling of the transmitted/received signals. However, to address the uncertainties and non-linearities imposed by channel equalization and hardware impairments, model-free ML techniques have become common in wireless communications [7, 8].

A. Model-based approaches

Channel estimation is an essential task for reliable communication [11]. However, it is more challenging in 6G architectures, wherein number of antennas in UM arrays is exceedingly huge [2, 11]. Furthermore, the multi-hop communications frameworks, such as vehicular networks [10, 13] and intelligent reflecting surface (IRS)-empowered systems [8, 14] make channel estimation even more demanding. In particular, vehicular networks have highly dynamic channels arising from rapid vehicular mobility (up to 150 km/h in 6G). Then, variation in weather conditions (resulting in a path loss of ~4 dB over 0.1-0.3 THz) causes frequent drop-outs and handovers [10, 12]. As a result, there exists an inherit uncertainty stemming from the dynamics of the wireless channel in both network architectures.

B. ML-based approaches

Massive MIMO channel estimation via ML is investigated in [7], where a convolutional-only neural network (CoNN)

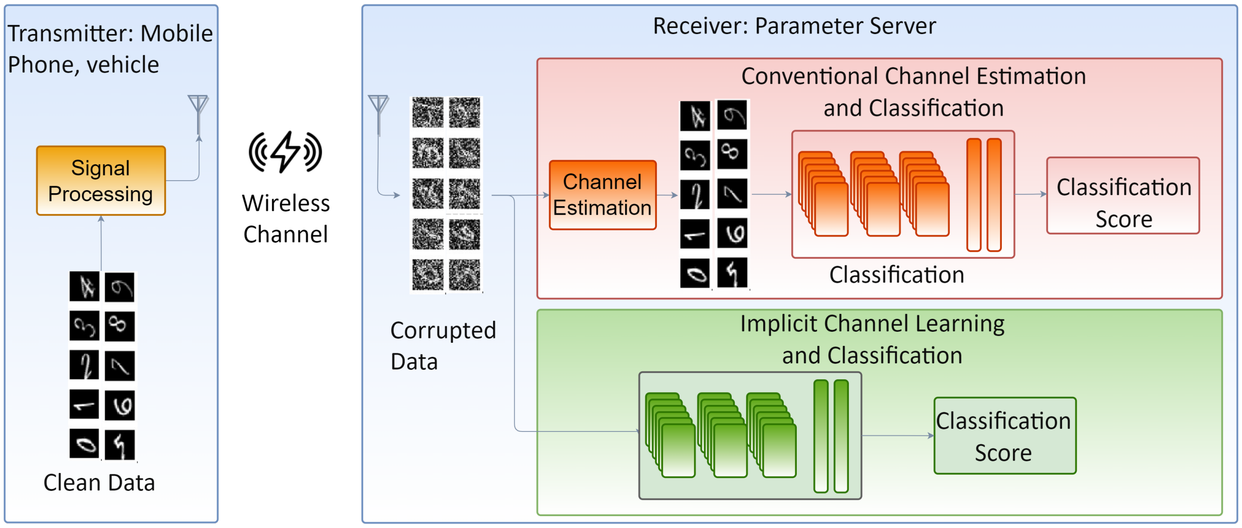

Fig. 3. Joint image classification and implicit wireless channel learning, wherein the clean data are processed and transmitted through the wireless channel to the PS. Next, the reconstructed channel-corrupted data are fed to the learning model for classification.

is designed. The input of the CoNN is the tentative channel matrix that is computed via least squares (LS) method and the output is the channel matrix from multiple subcarriers. The CoNN exhibits improved channel estimation accuracy in terms on the mean-squared error (MSE). But its input data requires matrix inversion, which is computationally inefficient for large antenna systems. To reduce the complexity in preparing the input data, [6] devised a CNN approach for channel estimation by feeding the CNN with the received pilot signals. This approach not only eliminates matrix inversion, but also estimates the full CSI (in the scale of the number of antennas) from reduced-rank input data (in the scale of number of pilot signals, fewer than the number of antennas). This is more applicable to massive MIMO systems, where low dimensional radio-frequency (RF) precoders limit the dimension of the received data [11]. Nevertheless, training in [7] and [6] is based on centralized schemes, for which the training data needs to be transmitted from the edge devices (mobile phones) to the PS. This results in huge communications overhead arising from large sizes of training datasets. This may be mitigated by employing federated learning (FL) [8, 10], wherein CNN parameters are computed at edge devices based on their local datasets not shared with the PS; instead, only the model parameters/updates are transmitted to the PS for model aggregation.

For vehicular networks, ML-based channel estimation in [13] used a multi-layer perceptron (MLP) architecture to enhance the tentative channel estimate (TCE) obtained via spectral-temporal averaging (STA). This incorporated time and frequency correlation ratios of two successive orthogonal frequency division multiplexing (OFDM) symbols. The input to MLP was STA-based TCE while the output was the true channel vector labels.

C. Shortcomings of ML-based channel estimation

The performance of ML-based channel estimation techniques is upper-bounded by the accuracy of the labeling algorithm, which is usually model-based [7, 13]. These algorithms generally rely on the specific mathematical model of the wireless channel and matrix processing steps such as singular value decomposition (SVD) [4], matrix inversion [7], and covariance computation [11]. As a result, labeling remains one of the major challenges in ML-based channel estimation. While physical layer applications such as beamformer design or resource allocation handle labeling via optimization techniques, the ML-based channel estimation methods [6, 7, 13] assume the label as true channel data. When imperfect channel data are designated as label, the estimation accuracy degrades in the MSE sense [8]. Obviously, the choice of label-generation algorithm is critical in ensuring the ML performance.

IV. Implicit Channel Learning in ML Tasks

Instead of designing dedicated learning models for both channel estimation and ML applications, both tasks could be performed jointly. Here, the ML model is trained on a training dataset while accounting for the imperfections or changes in the wireless channel. This allows the model to learn the corruptions in the ML data and perform the recognition tasks without performing channel estimation, leading to an overall reduction in the computational complexity and channel overhead.

The joint learning of the channel and the ML data requires preparing the training dataset for various channel conditions so that the ML model can extract the pattern in the input data and be robust against the corruptions of the wireless channel. Fig. 3 illustrates the processing chain of training data generation. Here, we generated three datasets each for both image and speech classification. The first `D1` was *clean* while the second (`D2`) contained the corrupted data. Third set (`D3`) is a mix of 50% clean and 50% corrupted data. During data generation, the wireless channel statistics were randomly changed for each image transmission and the signal-to-noise-ratio (SNR) was set to 15 dB. Each corrupted dataset included 100 distinct corrupted copies of the clean dataset to provide robustness

TABLE II

IMAGE (LEFT) AND SPEECH (RIGHT) CLASSIFICATION PERFORMANCE UNDER DIFFERENT CHANNEL CONDITIONS.

Learning Model: CNN with two convolutional layers (128@5×5 and 128@ 3×3) and a single fully connected layer (128 units)
Dataset: MNIST Handwritten Digits Dataset ($28\times28\times60,000$)
Classes: $\{0,1,2,3,4,5,6,7,8,9\}$

Training Data: D1 Test Data: Clean	Training Data: D1 Test Data: **Corrupted**	Training Data: D2 Test Data: Clean	Training Data: D2 Test Data: **Corrupted**	Training Data: D3 Test Data: Clean	Training Data: D3 Test Data: **Corrupted**
mmWave					
98.6%	**34.4%**	93.3%	**97.1%**	95.1%	**97.0%**
V2V - Rural LoS					
98.6%	**32.1%**	92.0%	**95.1%**	93.3%	**94.3%**
V2V - Urban LoS					
98.6%	**27.7%**	97.8%	**87.9%**	98.1%	**86.2%**
V2V - Urban nLoS					
98.6%	**15.2%**	97.1%	**87.7%**	98.0%	**85.8%**
V2V - Highway LoS					
98.7%	**30.4%**	97.9%	**90.3%**	98.3%	**88.4%**
V2V - Highway nLoS					
98.7%	**13.5%**	96.9%	**97.7%**	98.2%	**85.6%**
mmWave-V2V - Rural LoS					
98.3%	**27.5%**	90.4%	**92.2%**	91.1%	**91.9%**
mmWave-V2V - Urban LoS					
98.2%	**31.0%**	90.5%	**92.8%**	90.3%	**91.1%**
mmWave-V2V - Urban nLoS					
97.7%	**12.5%**	83.8%	**88.1%**	86.1%	**87.4%**
mmWave-V2V - Highway LoS					
98.4%	**32.0%**	90.4%	**90.3%**	89.0%	**89.8%**
mmWave-V2V - Highway nLoS					
97.1%	**17.6%**	84.4%	**88.1%**	86.5%	**87.2%**

Learning Model: CNN with four convolutional layers (16@5×5, 16@3×3, 16@3×3 and 16@3×3) and a single fully connected layer (128 units)
Dataset: Speech Command Dataset ($98\times50\times28,370$)
Classes: $\{\text{yes, no, up, down, left, right, on, off, stop, go}\}$

Training Data: D1 Test Data: Clean	Training Data: D1 Test Data: **Corrupted**	Training Data: D2 Test Data: Clean	Training Data: D2 Test Data: **Corrupted**	Training Data: D3 Test Data: Clean	Training Data: D3 Test Data: **Corrupted**
mmWave					
91.5%	**31.6%**	78.6%	**88.3%**	81.8%	**87.1%**
V2V - Rural LoS					
91.5%	**29.6%**	77.1%	**87.5%**	79.8%	**86.3%**
V2V - Urban LoS					
91.6%	**27.7%**	77.5%	**85.6%**	80.0%	**84.7%**
V2V - Urban nLoS					
91.5%	**15.2%**	76.7%	**85.5%**	78.4%	**84.7%**
V2V - Highway LoS					
91.5%	**27.2%**	79.0%	**87.0%**	80.2%	**86.4%**
V2V - Highway nLoS					
91.5%	**13.5%**	75.3%	**85.1%**	77.1%	**85.1%**
mmWave-V2V - Rural LoS					
91.5%	**27.3%**	78.7%	**85.1%**	82.6%	**84.0%**
mmWave-V2V - Urban LoS					
91.6%	**25.8%**	77.8%	**84.2%**	82.3%	**83.4%**
mmWave-V2V - Urban nLoS					
90.3%	**14.1%**	76.2%	**83.7%**	81.5%	**82.5%**
mmWave-V2V - Highway LoS					
91.4%	**26.3%**	78.2%	**86.5%**	82.5%	**85.1%**
mmWave-V2V - Highway nLoS					
91.4%	**12.4%**	72.1%	**83.4%**	81.8%	**82.0%**

against imperfections. As a result, the number of samples in these three training datasets were $600,000$, $6,000,000$ and $3,300,000$ ($283,700$, $2,837,000$ and $1,560,350$) for MNIST (Speech Command) dataset, respectively. For the MNIST (Speech Command) dataset, there were two test datasets, each of which had a size of $10,000$ ($4,000$) and they were generated separately from the training data.

Table II shows image and speech classification performance of the learning models for various channel conditions such as mmWave, V2V, and multi-hop mmWave-V2V channels. The V2V channel was tested for multiple delay profiles corresponding to different scenarios, such as Rural LoS, Urban LoS/nLoS (non-LoS), and Highway LoS/nLoS [12, 13]. We observe that the learning performance was poor if there was a mismatch between the training and test datasets. In particular, the model trained on D1 was unable to recognize the corrupted dataset, which includes the channel effects. On the other hand, the models trained on D2 and D2 were able to learn the corrupted ML data (image or speech) while implicitly learning the channel characteristics. Compared to the case conducted without channel estimation, they exhibit approximately 60% improvement for both image and speech classification. Furthermore, the model with D2 provided slightly higher accuracy than the one possessing D3 for the corrupted test data while the latter had a slight performance loss (approximately 1%) for clean test dataset. This is because the size of D3 is smaller, including both clean and corrupted data. Nevertheless, D3 presents satisfactory performance for both tasks with a smaller dataset. Consequently, this suggests that constructing the half of the dataset with channel effects can yield a reliable recognition performance as well as implicitly learning the channel effect on the ML data.

To compare the channel characteristics, the accuracy for mmWave-only channels degraded when combined with the V2V channel. This is explained by channel dynamics and the loss of beamforming gain, that mmWave usually leverages upon via multiple antennas. The reliability of ML tasks aggravated in multi-hop scenario, i.e., mmWave-V2V for all delay profiles because of error propagation when the inaccurately received ML data in one vehicle was transmitted to the BS via the mmWave channel. Among all delay profiles, those with nLoS propagation had the most severe conditions leading to low classification accuracy. In particular, the 'Highway nLoS' showed the worst performance while 'Rural LoS' fared the best for all tasks in both V2V and mmWave-V2V channels.

Comparing the classification performance of both ML tasks, higher (~7%) accuracy was obtained for image classification than speech recognition. Note that both had the same number of classes. This is obvious because the image dataset had more distinguishable patterns (e.g., handwritten digits) whereas the features in the spectrogram of the audio signals were less prominent. When both training and test datasets were corrupted, the performance improvement in both tasks was within ballpark of each other due to the similarity. That is to say, they

TABLE III
IMPLICIT LEARNING APPLICATIONS

Application	Loss function	Input	Output
Implicit channel learning	Cross-entropy cost for channel corrupted data	Channel corrupted data	Classification score
Implicit channel and beamformer learning	Cross-entropy cost for channel- and beamformer- corrupted data	Channel- and beamformer-corrupted data	Classification score
Implicit channel learning and semantic communication	Cross-entropy cost for channel corrupted data and mutual information	A sentence with semantic encoding	Recovered sentence
Implicit channel learning for beamformer design	MSE between label and predicted beamformers	Received pilot signals	Beamformer weights
Implicit channel learning for user localization	MSE between label and predicted locations	Received pilot signals	User locations/directions
Implicit channel learning for antenna selection	Cross-entropy cost for received pilots and antenna subarray configuration	Received pilot signals	Best antenna subarray index

both get close to the clean dataset performance (i.e., 98.6% and 91.5% for, respectively, image and speech).

V. Challenges and Future Research Directions

The vision of 6G embraces the deployment of ML models for various problems/scenarios at different layers of OSI model, e.g., in the application layer (image recognition) [9], physical to transport layer (resource allocation) [1, 10, 14], and physical layer (channel estimation, beamforming) [6, 7, 13]. Thus, it is a challenging task to employ these ML models for different applications simultaneously.

A. Learning-related challenges

Our implicit channel learning approach is one possible solution to reduce the complexity of ML-based methods by combining the channel estimation and recognition tasks. Likewise, similar multi-task learning techniques may be considered for future work by combining the channel learning and beamforming, localization, and antenna selection as detailed in Table III.

Unlike prior works on IRS-assisted communications [14] and integrated sensing and communications (ISAC) [15], which require *a priori* knowledge (e.g., user locations, see Table I), our approach is applicable to such emerging applications. The integration of implicit channel learning for semantic communications is another interesting avenue, wherein the transmission of "meaning" of the data is studied in an end-to-end learning manner [1].

B. Data-related challenges

The performance of implicit channel learning strongly depends on the data used to train the ML model. The training dataset should include various channel conditions to provide a sufficient representation for the wireless channel. Collecting the channel data in the field is a tedious process. In order to obtain the channel data under certain channel distributions, generative adversarial networks (GANs) are helpful [5]. In ML context, GANs are used as generative models to produce data that follow a certain target distribution. For example, GANs generate synthetic media in "deepfake" applications, where a person in an image or video is swapped with another person's likeness. In wireless communications, GANs are utilized to synthetically generate various channel conditions.

Elimination of channel acquisition in implicit channel learning applies only to the data-types related to ML tasks, which consume a large amount of wireless data traffic [9]. Thus, decision on performing the channel estimation may be based on the data-type. This information could be accessible from data tags in the physical layer. But then it leads to a resource allocation problem stemming from the assignment of data-type-based channel estimation operations.

C. Communications-related challenges

Although 6G enables the peak data rate of 100Gb/s, the transmission of training datasets to PS for model training still carries a significant overhead. The solution is 6G-enabled edge intelligence or FL, wherein the ML models are trained at the edge level [1, 10]. For certain physical layer applications (channel estimation and beamforming), FL has been shown [8] to provide ~15 times reduction in the communications overhead over centralized ML (CML) that collects the data from the edge to the PS for training. The FL-based architecture also keeps datasets at the edge, which is more privacy-preserving and communications-efficient. Nevertheless, FL-based approaches need further enhancements in terms of computational resources at the edge level. The possible future solutions may include sparsification, pruning, and quantization of the ML models.

VI. Summary and Concluding Remarks

We introduced implicit channel learning for ML applications in 6G networks. The proposed method jointly learns the features both in ML data and channel characteristics. By constructing a training dataset under the effect of various channel conditions in different propagation environments, the trained ML model becomes robust against the corruptions/imperfections.

Compared to model-based methods, the ML-based techniques enhance the estimation performance and robustness against the channel dynamics. This is particularly helpful for highly dynamic channels at mmWave and THz, which also employ extremely large arrays.

Our proposed method does not rely on true channel knowledge as training labels and provides an end-to-end learning framework. Thus, unlike existing ML-based channel estimation procedures, it is applicable to other physical layer applications, such as resource allocation, beamforming, and localization for diverse and emerging 6G scenarios, e.g., IRS-assisted networks, ISAC, and holographic communications.

The performance of the ML model relies on the training data collected under various channel conditions, for which the GAN-based methods are helpful to enrich the datasets. FL-based approaches are advantageous in terms of privacy and communications-efficiency. However, training very large learning models via FL could be difficult.

REFERENCES

[1] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6G: Vision, enabling technologies, and applications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 5–36, 2021.

[2] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, "Artificial-intelligence-enabled intelligent 6G networks," *IEEE Network*, vol. 34, no. 6, pp. 272–280, 2020.

[3] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, "Deep Learning for Wireless Communications: An Emerging Interdisciplinary Paradigm," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 133–139, 2020.

[4] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, "Deep learning coordinated beamforming for highly-mobile millimeter wave systems," *IEEE Access*, vol. 6, pp. 37 328–37 348, 2018.

[5] H. Ye, G. Y. Li, B.-H. F. Juang, and K. Sivanesan, "Channel agnostic end-to-end learning based communication systems with conditional GAN," in *IEEE Global Communications Conference Workshops*, 2018, pp. 1–5.

[6] A. M. Elbir, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "A family of deep learning architectures for channel estimation and hybrid beamforming in multi-carrier mm-Wave massive MIMO," *IEEE Trans. Cognit. Commun. Networking*, 2021, in press.

[7] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-based channel estimation for mmWave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 989–1000, 2019.

[8] A. M. Elbir, A. K. Papazafeiropoulos, and S. Chatzinotas, "Federated learning for physical layer design," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 81–87, 2021.

[9] R. Mayer and H.-A. Jacobsen, "Scalable deep learning on distributed infrastructures: Challenges, techniques, and tools," *ACM Comput. Surv.*, vol. 53, no. 1, pp. 1–37, 2020.

[10] J. Posner, L. Tseng, M. Aloqaily, and Y. Jararweh, "Federated learning in vehicular networks: Opportunities and solutions," *IEEE Network*, vol. 35, no. 2, pp. 152–159, 2021.

[11] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

[12] Z. Zhao, X. Cheng, M. Wen, B. Jiao, and C. Wang, "Channel estimation schemes for IEEE 802.11p standard," *IEEE Intell. Transp. Syst. Mag.*, vol. 5, no. 4, pp. 38–49, 2013.

[13] A. K. Gizzini, M. Chafii, A. Nimr, and G. Fettweis, "Deep learning based channel estimation schemes for IEEE 802.11p standard," *IEEE Access*, vol. 8, pp. 113 751–113 765, 2020.

[14] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1931–1945, 2021.

[15] W. Yuan, S. Li, Z. Wei, J. Yuan, and D. W. K. Ng, "Bypassing channel estimation for OTFS transmission: An integrated sensing and communication solution," in *IEEE Wireless Communications and Networking Conference Workshops*, 2021, pp. 1–5.