

Deep Learning for mmWave Beam-Management: State-of-the-Art, Opportunities and Challenges

Ke Ma, Zhaocheng Wang*, *Fellow, IEEE*, Wenqiang Tian,
Sheng Chen, *Fellow, IEEE*, Lajos Hanzo, *Life Fellow, IEEE*

Abstract—Benefiting from huge bandwidth resources, millimeter-wave (mmWave) communications provide one of the most promising technologies for next-generation wireless networks. To compensate for the high pathloss of mmWave signals, large-scale antenna arrays are required both at the base stations and user equipment to establish directional beamforming, where beam-management is adopted to acquire and track the optimal beam pair having the maximum received power. Naturally, narrow beams are required for achieving high beamforming gain, but they impose enormous training overhead and high sensitivity to blockages. As a remedy, deep learning (DL) may be harnessed for beam-management. First, the current state-of-the-art is reviewed, followed by the associated challenges and future research opportunities. We conclude by highlighting the associated DL design insights and novel beam-management mechanisms.

I. INTRODUCTION

According to Ericsson’s mobility report [1], the worldwide total monthly mobile data traffic will increase by 30% each year and reach 143 exabytes in 2026. To meet the ultra-high data traffic requirement, the enhanced Mobile Broadband (eMBB) mode of the fifth-generation (5G) wireless network has been designed for supporting high-speed access for users in hot-spot areas. Millimeter-wave (mmWave) communications, benefiting from abundant bandwidth resources spanning from 30GHz to 300GHz, have the potential of supporting Gigabits-per-second data rates.

However, mmWave carriers suffer from higher pathloss than those of conventional low-frequency communication systems. Fortunately, the short wavelength of mmWave signals allows more antennas to be integrated into both the base stations (BSs) and user equipment (UE). Therefore, large-scale antenna arrays can be used at the BS and UE sides to implement directional beamforming, so that the high pathloss can be compensated by the beamforming gain. To provide seamless high-quality services, beam-management has to be adopted to acquire and track the optimal BS and UE beam pair having the maximum received power.

However, these narrow beams impose beam-management challenges. On the one hand, numerous narrow candidate beams have to be created for covering the whole angular space, but this imposes substantial beam-training overhead. As

a further challenge, narrow beams are sensitive to blockages, making accurate beam-tracking more difficult.

Inspired by the stunning breakthroughs that deep learning (DL) has achieved in computer vision and natural language processing, DL has also been harnessed in wireless communications. Compared to mathematical model-based methods, DL enjoys a pair of key advantages. Firstly, mathematical tools generally rely on idealized assumptions, such as the presence of pure additive white Gaussian noise, which may not be consistent with practical scenarios. By contrast, DL adaptively learns the features of the channel in support of reliable beam-management. Secondly, the parameters of DL models capture the high-dimensional features of the propagation scenario, such as blockage locations and shapes in support of reliable beam-management, as detailed in the paper.

Based on these motivations, we survey the state-of-the-art by pursuing two avenues. Firstly, DL is utilized for extracting the nonlinear features inherent in the angular domain and time domain for implementing super-resolution beam-prediction and low-overhead beam-tracking along with predictive beam-switching [2]–[6]. Secondly, DL is adopted for learning the high-dimensional environmental features from side information and from the environmental feedback for supporting environment-specific beam-management [7]–[11]. Finally, future research challenges and opportunities are summarized, where we provide design insights in support of practical implementations, and envision novel beam-management mechanisms relying on DL.

II. WHY DL FOR BEAM-MANAGEMENT?

To illustrate DL’s superiority in beam-management, we compare DL to three conventional beam-management methods, as summarized in Table I at a glance.

Beam-search-based methods: Since the numbers of candidate beams at the BS and UE sides are usually limited, a straightforward technique of finding the optimal beam pair is exhaustive search, which sweeps all the candidate beam pairs and selects the best one. However, this method may impose excessive overhead. To address this problem, twin-level lower-overhead beam-search based on a hierarchical multi-resolution codebook is widely adopted [3], [4], where the first-level search aims for finding the optimal wide beam. Then the second-level search confirms the specific optimal narrow-beam direction within the limited range of the selected wide beam. Another low-overhead scheme is the interactive beam-search [4], where the candidate transmit and receive beams are swept separately. Nevertheless, these conventional beam-search-based methods do not model the intrinsic nonlinear relationships among the received signals of beam-training,

This work was supported in part by the National Key R&D Program of China under Grant 2018YFB1801102 and in part by the National Natural Science Foundation of China (Grant No. 61871253). This work was also supported by OPPO Research Fund.

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/P034284/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council’s Advanced Fellow Grant QuantCom (Grant No. 789028)

(Corresponding Author: Zhaocheng Wang)

TABLE I
COMPARISON BETWEEN DL AND CONVENTIONAL BEAM-MANAGEMENT METHODS.

Methods	Key features	Advantages	Disadvantages
Beam-search-based methods	Sweep candidate beam pairs for beam selection	Easy to implement	Considerable overhead
Model-based methods	Estimate optimal beam angles based on channel models	Exploit channel properties to reduce overhead	Rely on priori model assumptions
Conventional ML-based methods	Use conventional ML tools to deduce data-specific insight	Better performance relying on environmental adaptation	Limited fitting capability owing to relatively simple models
DL-based methods	Adopt DL to extract nonlinear features and adapt to wireless environments	<ul style="list-style-type: none"> • Adaptive feature extraction for reducing overhead • End-to-end learning without priori assumptions • Strong fitting capability by massive parameters 	High calculational complexity

leading to considerable overhead. By contrast, DL is able to extract these intrinsic nonlinear features to predict the optimal beam pair by only measuring a few beam pairs, so that the huge training overhead can be effectively reduced.

Model-based methods: Generally, the optimal beam pair is aligned with the angle-of-arrival (AoA) and angle-of-departure (AoD) of the strongest path. Therefore, finding the optimal beam pair can be formulated as an angle estimation problem according to the mathematical properties of the channel model. For example, compressed sensing has been broadly applied to estimate the AoA and AoD based on the angular sparsity of the mmWave channel [12]. Nevertheless, these model-based methods normally rely on priori assumptions, such as the channel's sparsity and quantization of the AoA/AoD. This makes their feasibility uncertain in practical scenarios. By contrast, DL is an end-to-end approach that does not require any model assumptions. Based on the back-propagation algorithm, the parameters of the DL models can be adaptively optimized in a data-driven manner for tackling the specific user scenarios.

Conventional machine learning (ML)-based methods: Usually, ML is used for deducing environment-specific insight from the collected data, where conventional ML-based methods mainly rely on relatively simple models. For example, considering the limited number of candidate beam pairs, finding the optimal beam pair can be formulated as a classification task, where the powerful support vector machine is widely adopted to implement the classifier [13]. Compared to these conventional ML tools, DL benefits from its multi-layer structure having a huge number of learnable parameters, and thus enjoys stronger fitting capability to address sophisticated beam-management problems in the face of environmental uncertainty.

To achieve higher beamforming gain and reduce beam-training overhead constitute fundamental targets of beam-management. The advantages of DL inspire two research routes for these targets. Firstly, DL can be utilized for accurately extracting the intrinsic nonlinear angular- and time-domain features for both beam-prediction and beam-tracking, hence reducing the training overhead while guaranteeing high beamforming gain. Secondly, DL can be harnessed to sense and adapt to the complex wireless environments for facilitating environment-specific low-overhead high-gain beam-management. In Sections III and IV, we will elaborate on the state-of-the-art concerning both research routes, as summarized in Table II.

Note that the dataset construction is one of the most

important issues for implementing these DL-assisted beam-management schemes. For supervised learning, the DL model optimizes its parameters based on the prediction loss of the training dataset. Specifically, each sample in the training dataset consists of the prediction input and the optimal beam label. To construct this dataset, the BS and UE usually perform beam-search to find the optimal beam pair as the label, and meanwhile they collect the corresponding prediction input. By contrast, for deep reinforcement learning (DRL), although no explicit training dataset is required, the DL model interacts with the wireless environment and leverages its feedback to optimize the model parameters. In this context, it is worth mentioning that Alkhateeb *et al.* generously provided a generic open-source DL dataset DeepMIMO [14]. Based on the environment construction and accurate ray-tracing, DeepMIMO succeeds in accurately modelling the real-world wireless features for facilitating DL-assisted beam-management studies.

III. LEARNING NONLINEARITIES FOR BEAM-PREDICTION AND BEAM-TRACKING

In practical scenarios, beam-management is faced with various sources of nonlinearities. DL is capable of accurately extracting the nonlinear features as the outputs at the intermediate layers of neural networks and exploiting them to facilitate efficient beam-management.

A. Super-Resolution Beam-Prediction

MmWave beams are usually generated by analog phase shifters, whose unit-amplitude constraint results in the power leakage, which is a nonlinear function of the pointing angle [4]. Briefly, as it transpires from Fig. 1(a), achieving high beamforming gain requires extremely accurate angular alignment. In case of misalignment, part of the beam's power leaks to side lobes, but as a benefit, this power leakage allows us to estimate the AoA and AoD of the dominant path by exploiting the signals received during beam-training. However, using conventional estimation methods to solve this problem is faced with two challenges. Firstly, this intrinsic nonlinear relationship is difficult to accurately model, and secondly, the complex multipath components may severely interfere with the angle estimation. To address these issues, DL may be used to accurately model the nonlinear power leakage and the multipath interference for estimating the AoA/AoD. Since the optimal beam direction is generally aligned with the AoA/AoD of the dominant path [2], [4], the beam prediction can be

TABLE II
DL FOR MMWAVE BEAM-MANAGEMENT: STATE-OF-THE-ART.

Learning nonlinearities for beam-prediction and beam-tracking				
Research routes	Refs.	Learning methods	DL models	Key features
Super-resolution beam-prediction	[2]	Supervised learning	CNN	Measure fixed angular-spacing beams for beam-prediction
	[3]	Supervised learning	CNN	Measure wide beams for narrow-beam prediction
	[4]	Supervised learning	LSTM	Measure partial high-SNR wide beams for narrow-beam prediction
Beam-tracking and predictive beam-switching	[5]	Supervised learning	LSTM	Integrate DL with conventional Bayesian estimation-based beam-tracking
	[6]	Reinforcement learning	DQN	Adjust beam-tracking range according to effective throughput rewards
	[3]	Supervised learning	LSTM	Predict the optimal beam in the middle of current and next beam-training instants without beam-search
Learning environments for beam-prediction and adaptive beam-management				
Research routes	Refs.	Learning methods	DL models	Key features
Side information-assisted beam-prediction	[7]	Supervised learning	FC-DNN	Use low-frequency CSI in co-located scenarios for mmWave beam-prediction
	[8]	Supervised learning	ResNet	Use vision images acquired at BS for beam-prediction
	[9]	Supervised learning	CNN	Use 3D panoramic scenes generated by environmental 3D point cloud and UE's location for beam-prediction
DRL-assisted adaptive beam-management	[10]	Reinforcement learning	Wolpertinger architecture	Design environment-specific beamforming codebooks based on beamforming gain feedback
	[11]	Reinforcement learning	DQN	Solve joint beam-management, power-control and interference-coordination, and use differential manners to reduce action space

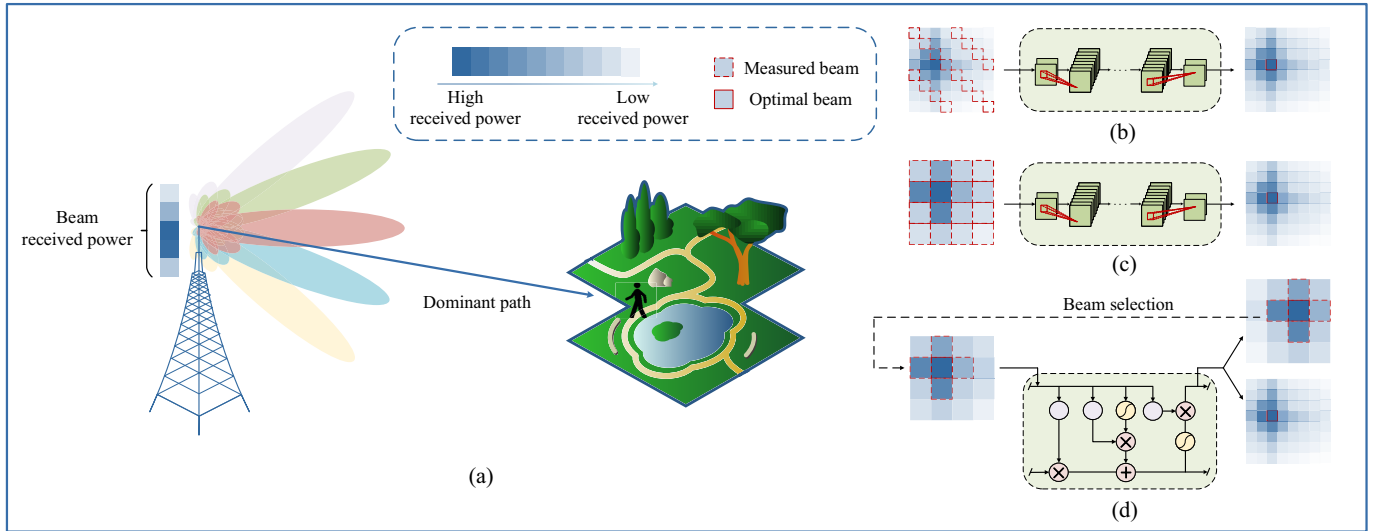


Fig. 1. (a) An example of the channel's power leakage; (b)~(d) illustration of various super-resolution beam-prediction schemes: (b) CNN-based prediction using fixed angular-spacing beams; (c) CNN-based prediction using wide beams; (d) LSTM-based prediction using partial high-SNR wide beams.

regarded as estimating the corresponding quantized AoA/AoD of the candidate beam directions, so that the optimal high-resolution beam direction can be predicted based on low-resolution beam-search for reducing the training overhead.

An intuitive technique of reducing the beam-training overhead is to only consider the specific beams having a fixed angular spacing and then use their received signals to predict the optimal beam [2], as shown in Fig. 1(b). Specifically, since the number of candidate beams is limited, this prediction is formulated as a classification task. A convolutional neural network (CNN) is used as the classifier, followed by a softmax function that converts the output into probabilities. Then the maximum-probability beam pair is selected. Supervised learning is adopted for the model optimization, which is comprised of the training stage and the prediction stage. At the

training stage, the BS and UE sweep all the candidate beams to find the optimal beam label, while the corresponding received signals of the fixed angular-spacing beams are collected as the prediction input, so that the training dataset is constructed. Then, the back-propagation algorithm is utilized for optimizing the model parameters based on the cross entropy loss of the training dataset. Once the model is well trained, supervised learning switches to the prediction stage. At this stage, only the fixed angular-spacing beams are measured to predict the optimal beam, so that the excessive training overhead can be effectively reduced. However, its performance may degrade due to the low signal-to-noise ratio (SNR) when the true AoA and AoD of the dominant path does not fall exactly in the middle of the main lobe of any candidate beam.

To fully cover the 360° angular space, Echigo *et al.* [3]

proposed a wide-beam-based optimal narrow-beam prediction scheme. As seen in Fig. 1(c), the wide and narrow beams can be naturally regarded as low-resolution and high-resolution beams. Consequently, Echigo *et al.* [3] compared super-resolution beam-prediction to a super-resolution image recovery problem, and adopted a CNN for the associated prediction.

Furthermore, as shown in Fig. 1(a), the leaked power of the beams which are angularly far from the AoA and AoD of the dominant path is low. Hence it is difficult to extract useful information from the corresponding received signals due to their low SNRs. Therefore, Ma *et al.* [4] proposed to consider only the subset of wide beams having high SNRs and use their received signals to predict the optimal narrow beam. This way the wide-beam-training overhead was further reduced. Specifically, a long-short term memory (LSTM) network was constructed for modelling the temporal AoA and AoD variations, as seen in Fig. 1(d). At the t -th time slot, the proposed network predicts not only the t -th optimal narrow beam, but also the $(t+1)$ -st optimal wide beam for selecting the neighboring high-SNR wide beams in the $(t+1)$ -st wide-beam-training.

Figure 2 compares the above DL-based schemes and the conventional noise-free beam estimation scheme based on wide beams in terms of their effective bandwidth efficiency, which takes the beam-training overhead into consideration. The dedicated frame structure is illustrated as follows. Beam-training is performed periodically, and the specific durations of the total communication session and of a single beam measurement are 1,000 ms and 0.1 ms [4]. The numbers of narrow candidate beams, fixed angular-spacing narrow beams [2], wide candidate beams [3] and high-SNR wide beams [4] are 64, 16, 16 and 5, respectively. The conventional scheme estimates the optimal narrow-beam direction based on the ratio of the beamforming gains between the optimal wide beam and the neighboring wide beams without considering the noise.

As expected, the bandwidth efficiency initially increases with the beam-training period owing to the reduced beam-training overhead. By contrast, long beam-training periods degrade the bandwidth efficiency, because the loss of beam-alignment is more likely to occur during longer data transmissions in mobile scenarios. It is clear that the DL-based schemes outperform their conventional counterparts, because DL models the nonlinear power leakage phenomenon more accurately. Furthermore, the partial-search-based high-SNR wide-beam-assisted scheme achieves higher bandwidth efficiency due to its lower training overhead, especially for short beam-training periods.

B. Beam-Tracking and Predictive Beam-Switching

Another nonlinearity in beam-management arises from the UE's mobility. As the UE moves closer to the BS, the angular variation of the dominant path would become faster as a function of their distance and vice versa. Moreover, the velocity and direction of UEs is time-variant. However, the conventional Kalman filter-based beam-tracking usually assumes a fixed AoA/AoD variation model, making it hard to

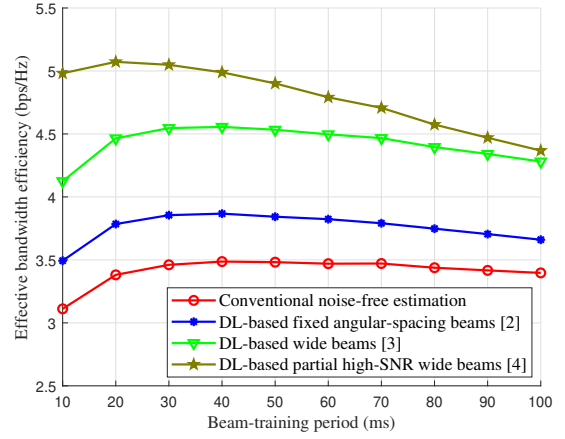


Fig. 2. Effective bandwidth efficiency as function of beam-training period for various schemes.

handle the nonlinear relationship between the UE's movement and angular variation, especially for high-speed scenarios. Therefore, DL is exploited to model these nonlinear variations for enhancing the beam-tracking accuracy attained.

Lim *et al.* [5] combined the conventional beam-tracking scheme with DL. At each beam-tracking action, firstly the LSTM network is exploited for fusing the previous channel state information (CSI) estimates and sensor measurements to extract the UE's movement features for predicting the *a-priori* AoA distribution. Based on this predicted distribution, the future beams are predicted and then measured for creating the more accurate *a-posteriori* estimate from the *a-priori* CSI estimate by using sequential Bayesian estimation. This scheme intrinsically integrates the expert knowledge of classical mathematical models with the adaptive fitting capability of DL, and thus achieves lower tracking errors than the conventional DL-based scheme that only relies on training data.

The UE's ever-changing mobility is another key issue influencing beam-tracking. Explicitly, when the velocity is high, the tracking range has to be increased and vice versa. Therefore, Zhang *et al.* [6] proposed to use DRL for seamlessly adapting to the tracking range. Concretely, the tracking action is jointly determined by the initial beam index and the size of the tracking beam subset. The proposed deep Q-network (DQN) interacts with the environment, adjusting the tracking range according to the effective throughput reward that takes both the beamforming gain and the tracking overhead into account.

As a further advance, a predictive beam-switching scheme based on DL was proposed in [3]. After each beam-training action, the LSTM network is utilized to predict the optimal beam in the middle of the current and next beam-training instants according to the previous received beam-training signals, which halves the training overhead. Echigo *et al.* [3] additionally integrated predictive beam-switching with super-resolution prediction for further reducing the training overhead.

IV. LEARNING ENVIRONMENTS FOR BEAM-PREDICTION AND ADAPTIVE BEAM-MANAGEMENT

In wireless environments, numerous scatterers having diverse locations, sizes and shapes appear, which form high-

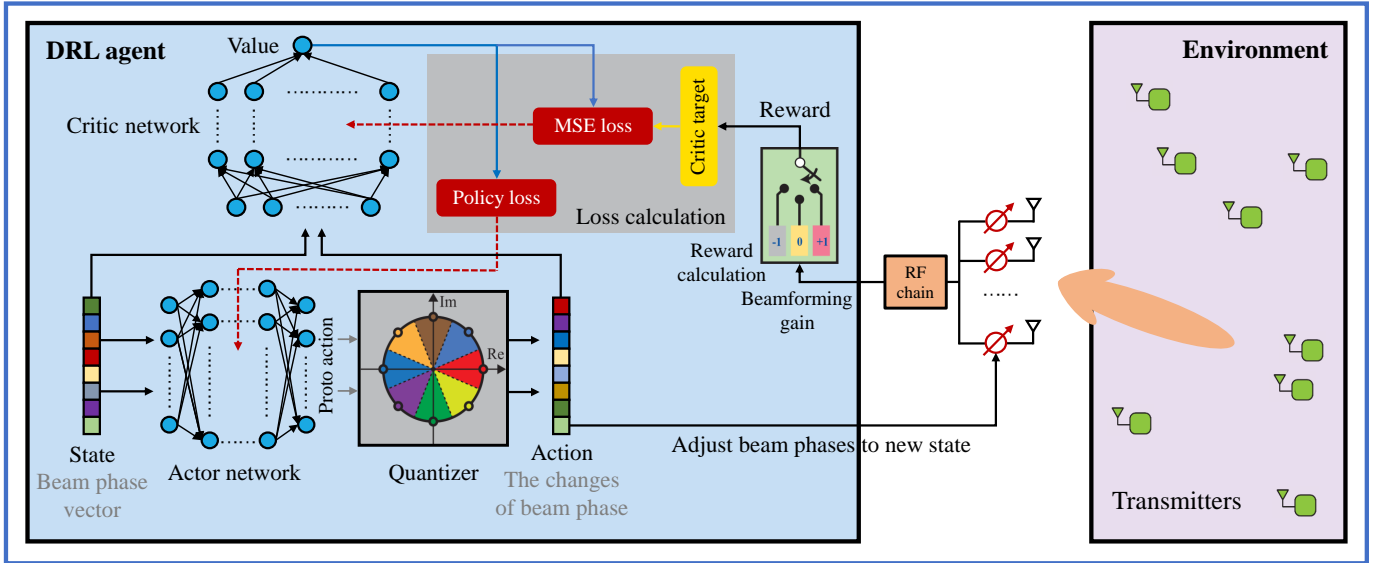


Fig. 3. DRL-based adaptive beam pattern design [10].

dimensional feature spaces. DL is capable of accurately extracting these complex scattering features captured at the intermediate layers of neural networks for facilitating beam-management.

A. Side Information-Assisted Beam-Prediction

When considering the big data scenarios of next-generation wireless networks, sophisticated inferences may be gleaned concerning the environmental features, which may be harnessed for mmWave beam-management. However, they are hardly ever leveraged by the conventional methods due to the lack of tractable mathematical models. By contrast, DL provides a promising tool for adaptively extracting these hidden environmental features for beam-prediction.

Usually, mmWave antennas are used also by the low-frequency BSs to reduce hardware cost. In this co-located scenario, the low-frequency and mmWave links perceive having similar propagation environments, and thus the mmWave channel enjoys analogous AoA and AoD features to the low-frequency counterpart. Therefore, DL may be exploited for revealing the complex relationships between the low-frequency and the mmWave channels by exploiting their shared environmental features. Alrabeiah *et al.* [7] conceived a fully connected-deep neural network (FC-DNN) for predicting the optimal mmWave beam based on low-frequency CSI, where the softmax function converts the output into probabilities. Based on these probabilities, Alrabeiah *et al.* [7] proposed to measure a limited subset of candidate beams having the highest probabilities, and to select the specific beam having the maximum received power as the optimal one, which achieved more accurate beam-alignment in exchange for a modest extra overhead.

On the other hand, vision images can provide surrounding environmental information for assisting beam-management. Therefore, Alrabeiah *et al.* [8] proposed to use images acquired from the cameras deployed at the BS to predict the optimal beam without channel measurements or beam-training. The well-known residual neural network (ResNet) in computer

vision is utilized to extract the image features. Besides, in order to model the propagation environments more comprehensively, a 3D scene-based beam-prediction framework was further investigated in [9]. At first, a 3D scene point cloud within the cellular coverage is constructed at the BS by massive images. Then, the BS combines the UE's location and the point cloud to generate the corresponding panoramic scene information, and a 3D-CNN is adopted to predict the optimal beam according to this information. At the expense of considerable overhead of constructing the point cloud offline, the 3D scene-based beam-prediction achieved higher accuracy than the conventional side information-assisted scheme.

B. DRL-Assisted Adaptive Beam-Management

In contrast to exploiting out-of-band side information, DRL constitutes an intelligent technique of sensing and adapting to the dynamic environments while only imposing negligible in-band overhead. The conventional learning-based methods usually use explicit channel knowledge for discovering environment-specific insights. By contrast, DRL integrates the interactive learning strategy of reinforcement learning with the strong fitting capability of DL, which could accurately model the high-dimensional environmental features and promptly adjust the beam-management policy, despite having limited environmental feedback.

The classical beamforming codebooks typically consist of numerous pre-defined beams without adaptation to the surrounding blockages and specific UE distributions, hence they impose excessive training overhead. To address this issue, Zhang *et al.* [10] proposed to apply DRL for designing the environment-specific beamforming codebooks. At first, all UEs are clustered into several groups based on the similarity of their channels. Then, each UE group is assigned with one DRL model to learn the corresponding best beam pattern, where the action is defined as the beam phase vector. However, the large number of antennas makes the action space very high-dimensional. Therefore, Zhang *et al.* [10] proposed a DRL framework based on the Wolpertinger architecture to narrow

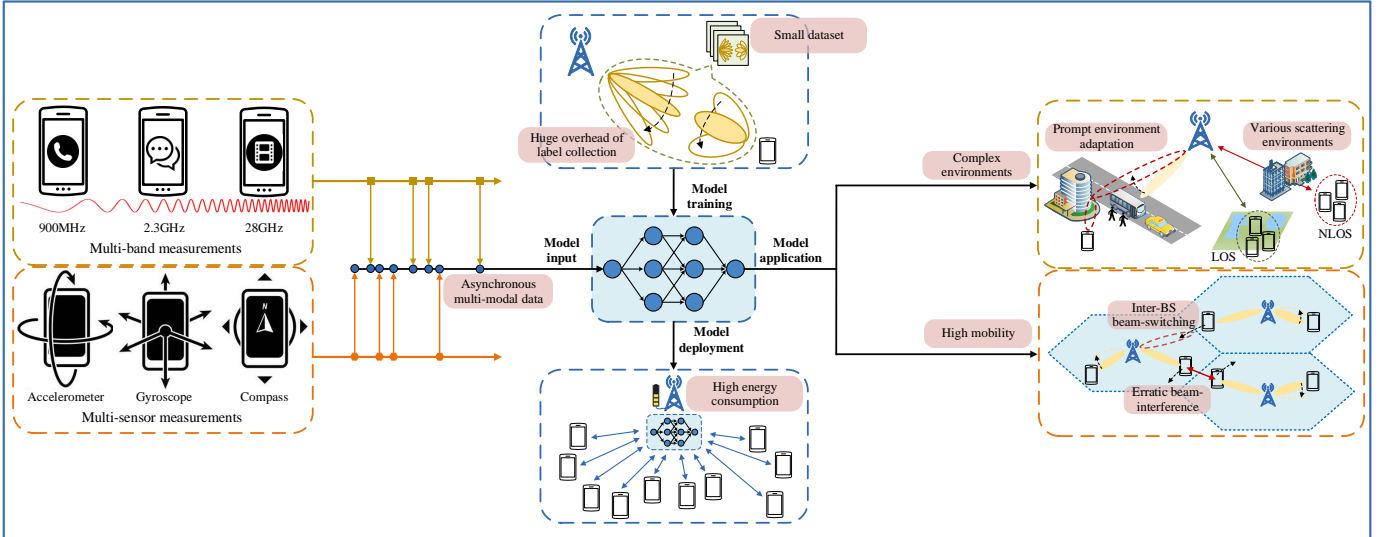


Fig. 4. Summary of research challenges for DL-assisted beam-management.

the size of the action space and avoid missing the optimal policy simultaneously, as shown in Fig. 3. Specifically, this architecture consists of an actor network and a critic network. The actor network interacts with the environment and generates the action, while the critic network evaluates the benefits of the action for optimizing both networks. At each learning step, the actor network firstly generates a continuous proto-action, which is quantized to discrete phases as the action. Then, the resultant phase vector is applied, and the reward is calculated according to the corresponding beamforming gain. Next, the mean-squared error (MSE) loss between the reward and the critic target learned from previous experience is calculated for optimizing the critic network. Finally, the action network is updated based on the policy loss assessed by the critic network. By iteratively optimizing the actor and the critic networks, the proposed model could accurately sense the environments according to low-overhead beamforming gain feedback, and ultimately achieved lower training overhead together with higher beamforming gain than the conventional codebooks.

For multiple BS scenarios, joint beam-management, power-control and interference-coordination is another crucial challenge due to its huge solution space. Therefore, Mismar *et al.* [11] utilized the DQN to adaptively sense the environment for tackling this high-dimensional problem. To ease the heavy burden of jointly optimizing the transmit powers and beams of numerous BSs, Mismar *et al.* [11] simplified each action as the difference from the previous value. Simulation results demonstrated that the proposed DQN model could approach the upper-bound performance.

V. RESEARCH CHALLENGES AND OPPORTUNITIES

Although DL has achieved beneficial performance gains in beam-management, there are still numerous open challenges for further study, as summarized in Fig. 4 at a glance.

A. Multi-Modal Integration Aided Beam-Management

Although a wide variety of side information has been leveraged to facilitate beam-management [7]–[9], the conception of a robust DL-based multi-modal integration framework

is in urgent demand for adaptively aggregating the features extracted from each input for accurate beam-alignment. As an initial attempt, Yang *et al.* [15] established a deep multi-modal learning framework for channel prediction, which relies on testing the performance of numerous models for guiding the design. To enhance the robustness, designing a generic model for handling any potentially missing modality constitutes one of the pivotal issues to be addressed. A feasible solution is to map the inputs into a joint latent feature space and then to reconstruct the missing modality via deep generative models. Next, tracking the temporal sequences of various inputs can facilitate robust beam-prediction, but their asynchronous nature imposes another crucial challenge. Neural ordinary differential equations provide an attractive continuous-time learning technique for accurately modelling these asynchronous sequences.

B. Efficient Training with Limited Dataset

Although the DL model trained by massive labeled data achieves satisfactory performance in beam-management, the huge overhead of collecting statistically relevant data is a challenge in practice. Therefore, few-shot learning, dedicated to utilizing a small number of samples for efficiently training the DL model, has great potential in beam-management. The conceptually simplest method of finding the optimal beam label is conventional beam-search, but again, this imposes an excessive overhead. To alleviate this burden, semi-supervised learning oriented data collection may be considered, where the optimal beam labels of only a few samples are gleaned, and then semi-supervised learning integrates the labeled data with a large amount of unlabeled data for improving the training performance.

C. Energy Consumption Concern

In practical scenarios, the beam-management of many UEs is performed simultaneously. However, given their high complexity, the DL-assisted beam-management methods may impose high energy dissipation. To tackle this problem, tailor-made light-weight DL models may be designed for beam-management relying on the adaptive tradeoff between per-

formance and complexity. Moreover, the application of high-efficiency parallel DL hardware and algorithms is also beneficial for reducing the energy consumption.

D. Robustness to Complex Environments

To provide seamless high-quality services, designing robust beam-management adaptive to complex propagation environments is crucial in practice. The existing DRL-based methods [10], [11] rely on limited environmental feedback for updating their model, hence their convergence may be relatively slow. Therefore, more prompt environment adaptation based on online DL is necessary for beam-management. On the other hand, the scattering environment has significant impact on the performance of DL-assisted beam-management. For example, compared to line-of-sight (LOS) scenarios, it is more challenging to accurately predict the optimal beam in non-line-of-sight (NLOS) scenarios. Consequently, a robust beam-management policy, which adopts DL for adjusting the beam measurement based on the specific environment for maintaining accurate beam-alignment, is worth investigating.

E. Nimble Beam-Management under High Mobility

The UE's high mobility typically leads to frequent BS-handover and beam-switching, hence requiring nimbler beam-management. In single-user scenarios, DL can be adopted for establishing an inter-BS beam-management framework, where BS selection and beam selection are performed jointly for attaining ultra-low handover latency. As for multi-user scenarios, the erratic nature of beam-interference constitutes a crucial challenge. Specifically, the mmWave beam impinging from a neighboring cell pointing towards the local UE imposes severe inter-cell interference. Considering frequent BS and beam variations due to high mobility, agile interference management is important for ensuring reliable services in dynamic mmWave networks. As a potential solution, the graph neural network (GNN) intrinsically matches the topology of wireless networks, which may be harnessed for handling dynamic inter-cell interference management.

VI. CONCLUSIONS

We commenced by elaborating on the motivation of applying DL in beam-management. Firstly, DL is eminently suitable for extracting the complex nonlinear features encountered. Secondly, the adaptive fitting capability of DL enables near-real-time fine-tuning required by the environmental fluctuations. Based on these motivations, DL was proposed for extracting both the intrinsic nonlinear angular- and time-domain features, hence inspiring super-resolution beam-prediction and beam-tracking along with predictive beam-switching. Furthermore, DL was utilized to sense and adapt to the propagation environments from side information and from the environmental feedback for designing environment-specific beam-management. Finally, we have provided DL design insights for the practical implementation of novel beam-management mechanisms.

REFERENCES

- [1] *Ericsson's mobility report*, Accessed: Nov. 1, 2021. [Online]. Available: <https://www.ericsson.com/en/reports-and-papers/mobility-report>.
- [2] W. Ma, *et al.*, "Machine learning for beam alignment in millimeter wave massive MIMO", *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 875–878, Jun. 2020.
- [3] H. Echigo, *et al.*, "A deep learning-based low overhead beam selection in mmWave communications," *IEEE Trans. Veh. Tech.*, vol. 70, no. 1, pp. 682–691, Jan. 2021.
- [4] K. Ma, *et al.*, "Deep learning assisted calibrated beam training for millimeter-wave communication systems," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6706–6721, Oct. 2021.
- [5] S. H. Lim, *et al.*, "Deep learning-based beam tracking for millimeter-wave communications under mobility," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7458–7469, Nov. 2021.
- [6] J. Zhang, *et al.*, "Intelligent interactive beam training for millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 2034–2048, Mar. 2021.
- [7] M. Alrabeiah and A. Alkhateeb, "Deep learning for mmWave beam and blockage prediction using sub-6 GHz channels," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5504–5518, Sep. 2020.
- [8] M. Alrabeiah, *et al.*, "Millimeter wave base stations with cameras: vision-aided beam and blockage prediction," in *Proc. VTC-Spring*, Antwerp, Belgium, May 2020, pp. 1–5.
- [9] W. Xu, *et al.*, "3D scene-based beam selection for mmWave communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1850–1854, Nov. 2020.
- [10] Y. Zhang, *et al.*, "Reinforcement learning of beam codebooks in millimeter wave and terahertz MIMO systems," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 904–919, Feb. 2022.
- [11] F. B. Mismar, *et al.*, "Deep reinforcement learning for 5G networks: joint beamforming, power control, and interference coordination," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1581–1592, Mar. 2020.
- [12] J. Choi, "Beam selection in mm-Wave multiuser MIMO systems using compressive sensing," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2936–2947, Aug. 2015.
- [13] Y. Long, *et al.*, "Data-driven-based analog beam selection for hybrid beamforming under mm-Wave channels," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 340–352, May 2018.
- [14] A. Alkhateeb, "DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications," in *Proc. ITA*, San Diego, USA, Feb. 2019, pp. 1–8.
- [15] Y. Yang, *et al.*, "Deep multimodal learning: merging sensory data for massive MIMO channel prediction," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1885–1898, Jul. 2021.

Ke Ma received his B.S. degree from Tsinghua University in 2019, where he is currently working toward the Ph.D. degree with the Department of Electronic Engineering, Tsinghua University. His current research interests include mmWave communications and intelligent communications.

Zhaocheng Wang (F'21) received his B.S., M.S., and Ph.D. degrees from Tsinghua University in 1991, 1993, and 1996, respectively. From 1996 to 1997, he was a Post Doctoral Fellow with Nanyang Technological University, Singapore. From 1997 to 2009, he was a Research Engineer/Senior Engineer with OKI Techno Centre Pte. Ltd., Singapore. From 1999 to 2009, he was a Senior Engineer/Principal Engineer with Sony Deutschland GmbH, Germany. Since 2009, he has been a Professor with Department of Electronic Engineering, Tsinghua University. He was a recipient of IEEE Scott Helt Memorial Award, IET Premium Award, IEEE ComSoc Asia-Pacific Outstanding Paper Award and IEEE ComSoc Leonard G. Abraham Prize.

Wenqiang Tian received his B.S. degree from Fudan University in 2010 and his Ph.D. degree from University of Chinese Academy of Sciences in 2015. Now, he is a senior Standardization Researcher of Guangdong OPPO Mobile Telecommunications Corp., Ltd. He has participated in the 5G standardization work and focused on physical layer design in 3GPP R15 and R16.

Sheng Chen (F'08) received his B.Eng. degree in control engineering from East China Petroleum Institute, China, in 1982, and his Ph.D. degree in control engineering from City University, U.K., in 1986, and the Doctor of Sciences (D.Sc.) degree from the University of Southampton, U.K., in 2005. From 1986 to 1999, He held research and academic appointments with the Universities of Sheffield, Edinburgh and Portsmouth, all in U.K. Since 1999, he has been with the School of Electronics and Computer Science, University of Southampton, U.K., where he holds the post of a Professor in intelligent systems and signal processing. He is a Fellow of the Royal Academy of Engineering (FREng), of the Asia-Pacific Artificial Intelligence Association and of the IET.

Lajos Hanzo (<http://www-mobile.ecs.soton.ac.uk>, https://en.wikipedia.org/wiki/Lajos_Hanzo) (F'04) received his Master degree and Doctorate in 1976 and 1983, respectively from the Technical University (TU) of Budapest. He was also awarded the Doctor of Sciences (D.Sc.) degree by the University of Southampton (2004) and Honorary Doctorates by the TU of Budapest (2009) and by the University of Edinburgh (2015). He is a Foreign Member of the Hungarian Academy of Sciences and a former Editor-in-Chief of the IEEE Press. He has served several terms as Governor of both IEEE ComSoc and of VTS. He is also a Fellow of the Royal Academy of Engineering (FREng), of the IET and of EURASIP. He is the recipient of the 2022 Eric Sumner Field Award.