Autoencoder for Position-Assisted Beam Prediction in mmWave ISAC Systems

Ahmad A. Aziz El-Banna, *Member, IEEE*, and Octavia A. Dobre, *Fellow, IEEE*

***Abstract*— Integrated sensing and communication and millimeter wave (mmWave) have emerged as pivotal technologies for 6G networks. However, the narrow nature of mmWave beams requires precise alignments that typically necessitate large training overhead. This overhead can be reduced by incorporating the position information with beam adjustments. This letter proposes a lightweight autorencoder (LAE) model that addresses the position-assisted beam prediction problem while significantly reducing computational complexity compared to the conventional baseline method, i.e., deep fully connected neural network. The proposed LAE is designed as a three-layer *undercomplete* network to exploit its dimensionality reduction capabilities and thereby mitigate the computational requirements of the trained model. Simulation results show that the proposed model achieves a similar beam prediction accuracy to the baseline with an 83% complexity reduction.**

***Index Terms*— Integrated sensing and communication (ISAC), millimeter wave (mmWave), autoencoder, beam prediction.**

I. Introduction

INTEGRATED sensing and communication (ISAC), a promising technology to be included in the 3GPP Release 19, is poised to be a key enabler for 6G systems [1], [2], [3], [4]. By jointly optimizing the spectrum and hardware resources, ISAC systems enable a realistic co-design of communication and sensing tasks [5]. Moreover, the millimeter-wave (mmWave) technology, offering significant improvements in spectral efficiency a nd s patial multiplexing capabilities, is essential for current and future wireless systems [6].

Specific t o t he m mWave t ransmission i s t he n arrow nature of the beams, which necessitates precise alignments in order to provide adequate coverage for users [7]. This technical limitation emphasizes the importance of robust beam prediction techniques in mmWave-based communication systems. Moreover, the substantial training overhead required for in-band channel estimation, particularly in highly mobile scenarios, poses a significant c hallenge f or b eam t racking. Nevertheless, given the imperative reliance of mmWave systems on line-of-sight (LoS) links, exploiting user position information, e.g., global positioning system (GPS), has emerged as a promising approach to reduce training overhead and facilitate efficient beam adjustments [6], [7].

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant RGPIN-2019-04123 and the Canada Research Chair program CRC-2022-00187. The associate editor coordinating the review of this letter and approving it for publication was T. Huynh-The. *(Corresponding author: Octavia A. Dobre.)*

Ahmad A. Aziz El-Banna is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1B 3X5, Canada, and also with the Faculty of Engineering at Shoubra, Benha University, Benha 13511, Egypt (e-mail: aaelbanna@mun.ca).

Octavia A. Dobre is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1B 3X5, Canada (e-mail: odobre@mun.ca).

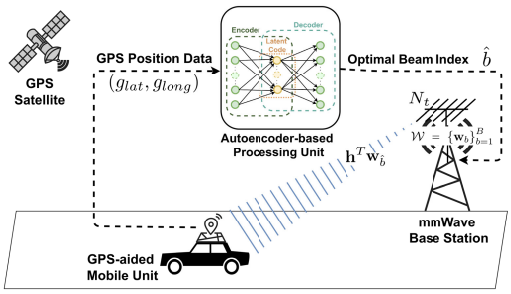

Fig. 1. A system model integrating the autoencoder beam prediction model with mmWave ISAC system.

Furthermore, in recent years, application of machine learning (ML) has been proven to provide an effective solution to numerous communication problems by enabling accurate modeling with reduced complexity compared to traditional statistical models [4], [8], [9]. The capabilities of ML techniques, in particular neural networks (NNs), in identifying the optimal beam for a sensing-aided mmWave system are investigated in [7]. While feed-forward NNs can effectively learn the complex mapping between the user position and the optimal beam, its computational complexity and memory footprint hinder its deployment on resource-constrained equipment. This highlights the need for more computationally efficient techniques, as real-time processing demands on modern devices continue to grow.

To fully leverage the capabilities of ML techniques for beam prediction, in this letter we propose a new lightweight autoencoder (LAE) architecture that significantly reduces the computational cost of state-of-the-art models, achieving approximately an 83% reduction in multiplications, additions, and model parameters. We design a computationally efficient LAE model, comprising only three hidden layers, and systematically select its configuration parameters to ensure robust generalization across multiple datasets that represent diverse environments of mmWave ISAC systems. The key contributions of this work are: (i) the development of a structured, low-complexity LAE tailored for beam prediction on resource-constrained devices; (ii) a comprehensive complexity analysis demonstrating substantial reductions in computational requirements; and (iii) empirical validation showing that the proposed model achieves prediction accuracy comparable to NNs across a wide range of real-world scenarios.

II. System Model and Problem Definition

A. System Model

We consider a mmWave ISAC system that utilizes an autoencoder for position-assisted beam alignment, as depicted in Fig. 1. A single-antenna mmWave mobile unit (MU), equipped with GPS, transmits its location information to an autoencoder-based processing unit which aids an N_t-antenna

stationary mmWave base station (BS) in selecting one of the B beamforming vectors, $\mathbf{w}_b \in \mathbb{C}^{N_t \times 1}$, from a codebook $\mathcal{W} = \{\mathbf{w}_b\}_{b=1}^{B}$, in order to transmit a complex symbol x. The received downlink signal can be expressed as $y = \mathbf{h}^T \mathbf{w}_b x + n$, where $\mathbf{h} \in \mathbb{C}^{N_t \times 1}$ represents the channel vector between the mmWave BS antennas and the mmWave MU antenna, and $n \sim \mathcal{N}_{\mathbb{C}}\left(0, \sigma^2\right)$ is the complex-valued Gaussian distributed noise with zero mean and variance σ^2.

B. Problem Definition

Considering the received signal power as the evaluation metric, the optimal beam maximizes the expected received power $P = \mathbb{E}\left[|y|^2\right]$ at the designated user. Consequently, the ideal beamformer $\mathbf{w}^\star$ is selected to receive the maximum power P at the BS, and the problem can be formulated as $\mathbf{w}^\star = \underset{\mathbf{w} \in \mathcal{W}}{\arg\max} \left|\mathbf{h}^T \mathbf{w}\right|^2$ [7]. The typical challenges associated with acquiring the channel information $\mathbf{h}$, and the reliance of mmWave systems on narrow, directional beams and LoS paths, motivate using real-time position information as an alternative to the direct channel estimation for the beam selection, i.e., bypassing the explicit estimation of $\mathbf{h}^T$ [7]. The problem now shifts to approximate $\mathbf{w}^\star$ by estimating $\hat{\mathbf{w}}$ that maximizes $\mathbb{P}\left(\hat{\mathbf{w}} = \mathbf{w}^\star | \mathbf{p}\right)$, where $\mathbf{p} = (g_{lat}, g_{long})$ denotes the 2-D position vector, with g_{lat} and g_{long} as the latitude and longitude GPS coordinates, respectively.

By collecting proper observations of the position and corresponding optimal beam pairs, ML techniques can be effectively utilized to learn the mapping between position and beam. Using a captured Q-labeled dataset, i.e., $\mathcal{Q} = \left\{\left(\mathbf{p}_q, \mathbf{w}_q^\star\right) : q = 1, \cdots, Q\right\}$, an ML algorithm can estimate the class probability vector $\mathcal{C} \in \{c_1, \cdots, c_B\}$, with $c_b = \mathbb{P}\left(\mathbf{w}_b = \mathbf{w}^\star\right)$ and $b \in \{1, \cdots, B\}$, in order to choose the beamformer with the highest probability, i.e., $\hat{\mathbf{w}} = \mathbf{w}_{\hat{b}}, \hat{b} = \underset{b \in \{1, \cdots, B\}}{\arg\max}\, c_b$.

III. Proposed Autoencoder (AE) Architecture

An AE is an NN architecture composed of two primary components: an encoder and a decoder [10]. Typically, the encoder compresses the input data into a compact latent code, while the decoder aims to reconstruct the original data from this latent code. The training objective is to minimize the reconstruction loss, which measures the discrepancy between the reconstructed and original data.

In this work, we utilize the dimensionality reduction properties of an *undercomplete* AE structure, wherein the code space has fewer dimensions than the input space, to learn a compact representation of the data that captures the most pertinent features, leading to a low-complexity model. Moreover, our proposed model is a lightweight AE named LAE, composed of three layers, with the middle (latent) layer shared between the encoder and decoder, as depicted in Fig. 1.

Given that the input features represent user positions (two correlated values) and the target is a single beam index, the beam prediction problem is inherently low-dimensional. To ensure minimal information loss during compression, a 2:1 ratio is chosen between the encoder/decoder layers and the latent layer, after evaluating both 2:1 and 4:1 ratios and determining that 2:1 provides better performance. To further balance complexity and accuracy, we conducted systematic experimentation and analysis across multiple autoencoder configurations, ultimately selecting two lightweight models, namely LAE-636 and LAE-424, that offer favorable trade-offs between computational complexity and predictive accuracy.

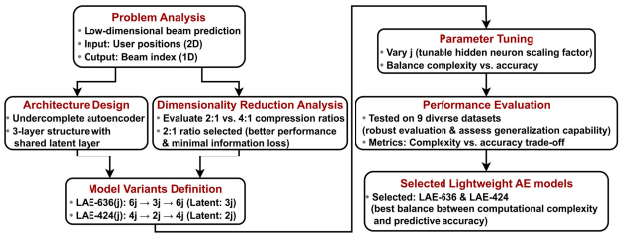

Fig. 2. LAE design and selection workflow.

Specifically, LAE-636(j) is a 3-layers AE that utilizes a 2:1 compression ratio and a latent dimension of $3j$, resulting in an architecture with $6j$, $3j$, and $6j$ hidden neurons in the first, latent, and last layers, respectively, where j is a predetermined number of hidden neurons. LAE-424(j) follows a similar structure but with a latent dimension of $2j$. As examples, LAE-636(10) is a basic AE with an initial layer consisting of 60 neurons, a latent space dimension of 30, and a final layer comprising 60 neurons, while LAE-424(10) is an AE with a 40-20-40 structure.

Finally, we trained and tested the models on nine datasets collected under diverse conditions, including variations in time of day, weather, and data volume. This demonstrates the robustness and generalization capability of the proposed LAEs while reducing the risk of overfitting to any specific scenario. The LAE design and selection workflow is shown in Fig. 2.

IV. Dataset Analysis and Best LAE Configurations Selection

A. Dataset Analysis

In this work, we utilize the dataset from [7], which is part of the comprehensive DeepSense dataset for 6G learning research [9]. All nine scenarios outlined in [7] are employed to conduct a thorough evaluation of the performance of the proposed LAE models in comparison to the baseline in [7], namely FCN3(256).[1] Two units were used to collect the dataset, emulating the mmWave ISAC system shown in Fig. 1. The first unit, a stationary mmWave BS equipped with a 16-element 60 GHz-band phased array, outputs a 64-element vector, along with the received power at each beam.[2] The second unit, a mobile vehicle, represents the MU and is

[1] We select FCN3(256), a three-layer NN with 256 hidden neurons, as a baseline due to its superior performance within the evaluated ML models [7].

[2] The first unit includes additional hardware like cameras and radar, but this work focuses on position-to-beam mapping using beam power data and GPS coordinates. The BS uses analog beamforming with a 16-element phased array, selecting one beam from a 64-beam codebook per user. The dataset is single-user, as per the testbed design. Details on the hardware and system setup are in [11]. Extending this framework to multi-user beam prediction is a promising future direction requiring additional data and system-level considerations. Moreover, future work could explore sensor fusion approaches, such as integrating GPS with inertial, radar, or vision sensors, to improve localization robustness and beam prediction accuracy under varying localization precision and environmental conditions [9], [13].

TABLE I

FEATURES OF THE NINE DATASET SCENARIOS FOR STREET-LEVEL MMWAVE COMMUNICATION [9], [11]

Category Description	Scenarios	# Samples
Day-time Scenarios	1	2,411
	3	1,487
	6	915
	7	854
	8	4,043
	9	5,964
Night-time Scenarios	2	2,974
	4	1,867
Night-time Scenario, Rainy Weather Condition	5	2,300

equipped with a 60 GHz omni-directional mmWave transmitter and a GPS receiver capable of outputting 10 location measurements per second and supports real-time kinematic (RTK) for centimeter-level accuracy. However, since the exact GPS mode used during data collection is not specified, we conservatively consider a typical differential GPS accuracy ($\sim$2.5 m) as the upper bound during model training. However, it worth mentioning that the proposed model is trained on real-world data collected under these conditions, allowing it to implicitly learn and adapt to the inherent imperfections in positioning accuracy. Finally, the dataset was collected from multiple locations across the Arizona State University campus, capturing a diverse range of real-world scenarios.

A summary of the features and the number of data samples collected for each scenario is presented in Table I. As can be observed in this table, the nine scenarios encompass a diverse range of operating conditions, with each scenario comprising hundreds to thousands of data samples. Scenarios 5, 6, and 7 present the most challenging conditions for training ML models. The dataset for scenarios 6 and 7 contains a limited number of captured samples, which can potentially constrain the learning capabilities of ML algorithms, particularly those whose performance is significantly influenced by the size of the training dataset, such as NNs. Further, the data collected for scenario 5 was acquired under rainy weather conditions, which can degrade the mmWave signal propagation due to increased absorption and scattering losses. This can lead to noisy data, making it more difficult for the ML algorithms to learn the underlying position-beam mapping function.

Finally, it is important to note that although no dataset can fully represent all possible deployment scenarios, by training and testing the model across this diverse set of conditions, we ensure that the learned representations are not tailored to any single data distribution. Instead, the model learns to capture underlying spatial and propagation patterns that are common across environments, resulting in a generic and deployment-ready solution.

B. Best LAE Configurations Selection

1) Configuration (1): LAE-636: In this analysis, we examine the LAE-636 structure with a range of hidden neuron values $j = \{20, 21, \ldots, 30\}$[3] and compare the achieved performance, i.e., top-1 beam prediction accuracy, of each structure to the baseline FCN3(256). Figure 3 (left plot) presents a heatmap illustrating the accuracy differences between the baseline and the tested LAE-636 structures for each of the nine scenarios described above.

In addition, we employed the *minimum-maximum difference (MinMaxDiff)* evaluation metric to determine the largest accuracy difference between the baseline and LAE structures for each scenario. Besides, the *minimum-average difference (MinAvgDiff)* metric was used to further differentiate the models by calculating the average of the accuracy differences between the LAE structures and the baseline across all the entire nine scenarios. The former metric identifies the model with the largest accuracy drop, while the latter provides an overall assessment of the model's performance relative to the baseline, taking into account both gains and losses across all testing scenarios. The *MinAvgDiff* metric also considers the individual accuracy gains of the LAE structures in specific scenarios. For instance, LAE-636(30) exhibits accuracy gains of 1.32%, 0.33%, and 0.74% over the baseline in scenarios 3, 7, and 9, respectively, demonstrating performance advantages not evident in the *MinMaxDiff* metric.

Finally, LAE-636(30) is selected for further analysis due to its optimal performance in terms of both *MinMaxDiff* and *MinAvgDiff*, as shown in Fig. 3 (right plot).

2) Configuration (2): LAE-424: Similarly, the performance of the LAE-424 structure was assessed for a range of hidden neurons $j = \{30, 31, \ldots, 40\}$. The achieved top-1 beam prediction accuracy of each structure was compared to the baseline FCN3(256) model. Figure 3 (middle plot) provides a detailed comparison of the accuracy differences between the examined LAE-424 structures and the baseline for each scenario. The LAE-424(38) and LAE-424(32) structures are selected as they exhibit the lowest *MinMaxDiff* and *MinAvgDiff*, respectively, as shown in Fig. 3 (right plot).

In the analysis of Fig. 3, hyperparameter tuning is performed to optimize the LAE models. The following hyperparameter spaces are explored: batch size (BZ) $= \{4, 8, 16, 32, 64, 128\}$, learning rate (LR) $= \{0.0001, 0.001, 0.01\}$, and activation function (AF) $= \{Sigmoid, ReLU\}$. The optimal hyperparameter values are found to be $BZ = 16$, $LR = 0.01$, and $AF = ReLU$.[4]

V. RESULTS AND DISCUSSION

A. Numerical Results

In this subsection, we analyze the top-1 and top-k beam prediction accuracies to evaluate the performance of the proposed LAE models in comparison to the baseline. The large number of employed beams (64) transforms the beam prediction problem into a challenging multi-label classification task. As such, to provide a more lenient measure of performance and assess model robustness, we include both

[3]We tested other neuron structures outside the range of 20-30 for AE-636 and 30-40 for AE-424. However, these structures either performed worse or similarly to those within the specified ranges, while also being more complex. Therefore, we focused our analysis on the 20-30 range for AE-636 and 30-40 range for AE-424.

[4]Notably, the main framework code and the datasets employed in this study are publicly accessible online [7], [11].

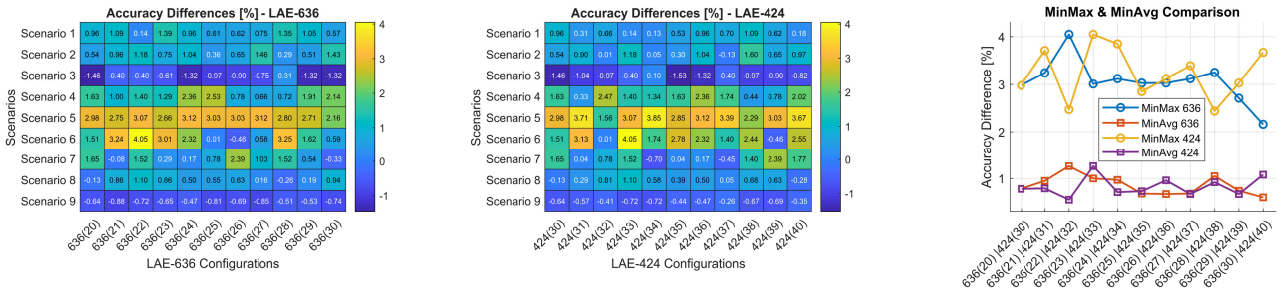

Fig. 3. Accuracy differences [%] between the baseline, i.e., FCN3(256), and configuration (1), i.e., LAE-636, (left) and configuration (2), i.e., LAE-424, (center). *MinMaxDiff* and *MinAvgDiff* values for both LAE-636 and LAE-424 configurations (right).

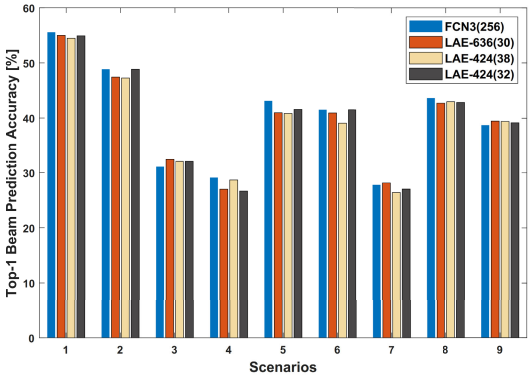

Fig. 4. Top-1 beam prediction accuracy for the LAE and baseline models.

top-1 and top-k[5] beam prediction accuracy as performance indicators. In addition to the analysis of beam prediction accuracy, we analyze the power loss between the predicted and actual beams, and the effect of changing the codebook size on the model performance.

Utilizing the aforementioned dataset, we optimize the LAE structures to achieve performance comparable to that of the baseline across the entire range of the real-world scenarios. Figure 4 shows the top-1 beam prediction accuracy of the pre-selected LAE models, i.e., LAE-636(30), LAE-424(38), and LAE-424(32), compared to the FCN3(256) baseline, for all scenarios. As depicted in Fig. 4, the selected LAE models perform similarly to the baseline across all nine scenarios. Moreover, Fig. 3 demonstrates that the three LAE models achieve a minimal accuracy gap of up to 2.47% relative to the baseline.

Additionally, Fig. 5 provides a comparative analysis of the top-k beam accuracy for the different models, with a focus on scenario 5 (rainy conditions), scenarios 3, 6, and 7 (small datasets), and scenario 9 (large dataset). As depicted in Fig. 5, the LAE and the baseline models exhibit similar generalization patterns, emphasizing the comparable performance achieved by the proposed LAE architecture. We select the LAE-424(32) model for the upcoming analysis as it attains the closest performance to the baseline in terms of both top-1 and top-k beam prediction accuracy.

Further, a rigorous metric for assessing system performance is the power loss between predicted and actual beam; the average power loss L can be expressed as $L_{[dB]} = 10\log_{10}\left(\frac{1}{Q}\sum_{q=1}^{Q}\frac{P^q_{\mathbf{w}^\star}-P_n}{P^q_{\hat{\mathbf{w}}}-P_n}\right)$, where $P^q_{\mathbf{w}^\star}$ and $P^q_{\hat{\mathbf{w}}}$ are the ground truth beam and predicted beam powers, respectively, and P_n is the scenario noise power [7]. Table II presents a comparative analysis of the power loss for both the selected LAE-424(32) model and the baseline model. As inferred from the table, the proposed LAE-424(32) model exhibits comparable performance to the baseline across the entire range of scenarios and it further outperforms the baseline in five scenarios, i.e., scenarios 1, 4, 6, 7, and 9.

Finally, we evaluate the top-1 beam prediction accuracy of the proposed LAE-424(32) and baseline models for different codebook sizes in Fig. 6. The figure shows that the proposed LAE-424(32) model acquires similar performance to the baseline across different numbers of employed beams. Notably, the proposed LAE-424(32) model exhibits superior performance to the baseline for a greater number of scenarios when smaller codebook sizes are employed.

B. Complexity Evaluation

An NN-based model constructed from a set of fully-connected dense layers S requires a total number of parameters equal to $\theta_D = \sum_{i=1}^{S}\theta_{di}$, with $\theta_{di} = m^{hd}_{di}\left(2m^{hd}_{d(i-1)}+1\right)$, where m^{hd}_{di} is the number of hidden neurons in layer i and $m^{hd}_{d(i-1)}$ is the number of hidden neurons in its previous layer $i-1$ [12]. Similarly, the number of required multiplications (MUL) or additions (ADD) can be computed as $\sum_{i=1}^{S} m^{hd}_{di} m^{hd}_{d(i-1)}$ [12].[6]

Table III presents the complexity reduction achieved by the candidate LAE models compared to the baseline. As can be seen from the table, the highest complexity reduction of over 83% in terms of the number of multiplications, additions, and memory requirements, i.e., model parameters, is offered by the LAE-424(32) model. The remaining LAE models also demonstrate significant complexity reduction, ranging from approximately 70% to 77%.

Eventually, the proposed LAE-424(32) model offers a significant complexity reduction of over 83% compared to the baseline, while maintaining a high level of performance with less than 2.5% loss in beam prediction accuracy as evident from Table III. Furthermore, LAE-424(32) has an average

[5]Regardless of the primary objective of identifying the most effective beam, i.e., the top-1 predicted beam, the selection of subsequent beams can still contribute to covering the designated user, although they are not the optimal choice.

[6]Both the FCN3(256) and AE models employ a ReLU activation function, which is assumed to require a single floating-point operation [12]. Consequently, the computational cost associated with ReLU activations is neglected in our calculations.

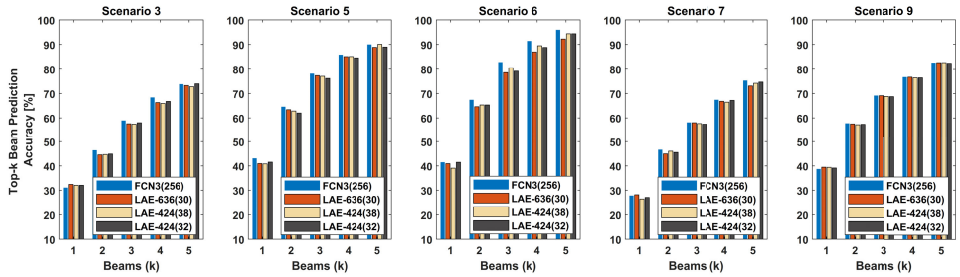

Fig. 5. Top-k beam prediction accuracy for scenarios 3, 5, 6, 7 and 9.

TABLE II
POWER LOSS FOR THE PROPOSED LAE-424(32) AND BASELINE MODELS

$L_{[dB]}$	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9
FCN3(256)	0.27	0.43	2.05	2.46	1.04	0.36	1.56	1.22	2.63
LAE-424(32)	0.26	0.45	2.32	2.44	1.14	0.33	1.46	1.25	2.6

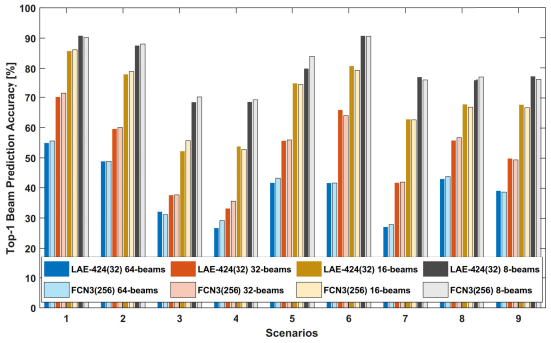

Fig. 6. Top-1 beam prediction accuracy at different codebook sizes (number of beams) for the proposed LAE and baseline models.

TABLE III
QUANTITATIVE COMPARISON IN TERMS OF COMPLEXITY AND ACCURACY

Criteria	Model	LAE-636(30)	LAE-424(38)	LAE-424(32)
Complexity Reduction	MUL/ADD [%]	70.07	77.61	**83.22**
	Parameters [%]	69.90	77.43	**83.05**
Accuracy Difference	Maximum [%]	**2.16**	2.44	2.47
	Average [%]	0.60	0.92	**0.55**

accuracy degradation of only 0.55% compared to the baseline across all tested scenarios, as indicated in Table III, which demonstrates the model's ability to generalize to different environmental conditions, highlighting its robustness and real-world applicability.

In summary, this work introduced an optimized LAE architecture for beam prediction in mmWave ISAC systems, with the LAE-424(32) model emerging as a particularly effective configuration. Through systematic structural tuning and parameter selection, the model achieves accuracy on par with the baseline while delivering substantial reductions in computational and memory demands, demonstrating the effectiveness of architectural optimization in enabling lightweight and practical deployment.

VI. CONCLUSION

In this letter, we proposed a lightweight autoencoder (LAE) model for position-assisted beam prediction. The proposed LAE model, constructed from an *undercomplete* structure with only three layers, leverages dimensionality reduction to efficiently learn the mapping between position and beam. Evaluated on a comprehensive real-world dataset encompassing nine distinct mmWave ISAC communication scenarios, the proposed model demonstrates comparable performance to the baseline, achieving an average beam prediction accuracy loss of only 0.55% while significantly reducing computational complexity by over 83%.

REFERENCES

[1] F. Liu et al., "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[2] S. Lu et al., "Integrated sensing and communications: Recent advances and ten open challenges," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19094–19120, Jun. 2024.

[3] C. Qi, W. Ci, J. Zhang, and X. You, "Hybrid beamforming for millimeter wave MIMO integrated sensing and communications," *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1136–1140, May 2022.

[4] U. Demirhan and A. Alkhateeb, "Integrated sensing and communication for 6G: Ten key machine learning roles," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 113–119, May 2023.

[5] F. Liu et al., "Seventy years of radar and communications: The road from separation to integration," *IEEE Signal Process. Mag.*, vol. 40, no. 5, pp. 106–121, Jul. 2023.

[6] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges," *Wireless Netw.*, vol. 21, no. 8, pp. 2657–2676, Apr. 2015.

[7] J. Morais, A. Bchboodi, H. Pezeshki, and A. Alkhateeb, "Position-aided beam prediction in the real world: How useful GPS locations actually are?" in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2023, pp. 1824–1829.

[8] M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Machine learning-based self-interference cancellation for full-duplex radio: Approaches, open challenges, and future research directions," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 21–47, 2023.

[9] A. Alkhateeb et al., "DeepSense 6G: A large-scale real-world multi-modal sensing and communication dataset," *IEEE Commun. Mag.*, vol. 61, no. 9, pp. 122–128, Sep. 2023.

[10] P. Baldi, "Autoencoders, unsupervised learning, and deep architectures," in *Proc. ICML Workshop Unsupervised Transf. Learn.*, 2012, pp. 37–49.

[11] Wireless Intelligence Lab. (2024). *DeepSense Testbeds*. Wireless Intelligence Lab. DeepSense6G. [Online]. Available: https://www.deepsense6g.net/data-collection/

[12] M. Elsayed, A. A. A. El-Banna, O. A. Dobre, W. Shiu, and P. Wang, "Hybrid-layers neural network architectures for modeling the self-interference in full-duplex systems," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6291–6307, Jun. 2022.

[13] B. Salehi, D. Roy, T. Jian, C. Dick, S. Ioannidis, and K. Chowdhury, "$Omni - -CNN$: A modality-agnostic neural network for mmWave beam selection," *IEEE Trans. Veh. Technol.*, vol. 73, no. 6, pp. 8169–8183, Jun. 2024.