

PERFORMANCE EVALUATION OF POSITION-AIDED BEAM PREDICTION USING EMPIRICAL DATA AND ANN IN 5G NR V2I COMMUNICATIONS

VINODH KUMAR MINCHULA*, UDVITHA G,
SUPRAJA REDDY A, SHARATHCHAND K

Chaitanya Bharathi Institute of Technology, Hyderabad, India

*Corresponding Author: vinodhkumarm_ece@cbit.ac.in

Abstract

Millimeter-wave systems depend on narrow beams for attaining required receive signal power, but this requires larger beam training overhead predominantly it's crucial in high mobility environments. Thereby, the proposed beam prediction approach mitigates this need of higher training overhead by leveraging user location data. This research investigates the performance of position-aided beam prediction using ANN models trained on real-world V2I datasets, leveraging large-scale empirical data to evaluate and compare CNN, RNN, and LSTM architectures, thereby moving beyond prior studies that rely primarily on synthetic or simulation-based datasets. Among the models, LSTM achieves the best performance, consistently reaching 38.04% for Top-1, 56.83% for Top-2, 69.5% for Top-3, 77.74% for Top-4, 82.96% for Top-5 beam prediction accuracies across 22,814 samples in nine scenarios, outperforming RNN by 4.24% and CNN by 45.51% making it the most effective solution for real-world 5G NR mmWave beam alignment in V2I Communications. LSTM achieves higher mean value approximately 20.3% in Top-1 accuracy, 32.8% in Top-5 accuracy than CNN and smaller consistent gain of around 2% at Top-1 & Top-5 than RNN. Heatmaps also reveal that CNN produces noisy outputs and RNN provides stable and consistent results with evenly streaked outputs, while LSTM delivers robust, scalable, and accurate predictions, making it the most effective solution for real-world mmWave beam alignment.

Keywords: 5G NR, Beam prediction, CNN, mmWave, MIMO, LSTM, RNN, V2I.

1. Introduction

Due to severe path loss and susceptibility to blockage in mmWave channels, 5G NR relies on narrow high-gain directional beams. Beam prediction plays a critical role in adaptive beamforming, where the goal is to dynamically select the optimal transmit-receive beam pairs particularly in high-frequency mmWave bands of 5G NR. Beam prediction minimizes the beam training overhead, thereby reducing latency, computational complexity and improving link robustness, spectral efficiency, and throughput. It becomes indispensable in high-mobility scenarios—such as vehicular networks or dense urban deployments—where rapid channel variation demands real-time, low-latency beam alignment.

Ultimately, this capability is foundational for enabling URLLC and eMBB in advanced 5G use cases like V2X. GPS-based position data can suggestively enhance beam prediction in real-world 5G V2I scenarios; however, its efficiency is restricted by aspects such as location inexactitudes and dynamic environments, highlighting the need for robust, data-driven models that can adjust to practical positioning challenges [1].

Deep learning techniques offer a powerful solution for beam prediction in 5G systems by assisting accurate, real-time beam selection, eventually improving communication efficiency, and system performance in dense environments [2]. A deep learning-based methodology facilitates quick and precise beam prediction in 5G and 6G networks, significantly minimizing communicating overhead and improving overall network efficiency [3]. Neural network-based beam management in mmWave massive MIMO systems [4-6] develops multi-user beam selection accurateness and system throughput while reducing computational complexity.

In GPS-assisted mmWave V2I communication, larger beam codebook sizes decrease the accuracy of machine learning-based beam prediction, while dataset split percentages have negligible impact—signifying that multimodal sensing could increase performance, nevertheless at the cost of added system complexity [7-9]. Leveraging the spatial correlation between sub-6 GHz and mmWave channels, the SPBPNNet deep learning model presents more effective mmWave beam prediction, reducing pilot overhead and upgrading performance in various environments [10].

An innovative beam prediction system for future 6G networks combines Convolutional Neural Networks (CNNs) and Artificial Neural Networks (ANNs), signifying their potential to enhance beamforming and signal quality in next-generation wireless communications. Efficiently adjusted deep learning models can more accurately estimate future signal blockages and identify optimal communication beams in mmWave networks, resulting in greater network performance and more reliable connections in dynamic environments.

Neural networks embrace significant promise for improving millimeter-wave beamforming in future wireless systems like 5G and beyond, as they can make beam alignment and channel estimation more accurate and efficient, eventually leading to upgraded network performance [11-15]. Vision-aided mmWave beam tracking in [16], shows that camera-based sensing can be used for wireless link adaptation. The work in [17] established a fundamental multimodal framework that continues to impact mmWave AI research by investigating deep learning-driven blocking and beam prediction using sub-6 GHz channels.

However, a critical research gap exists: the majority of these prior studies, rely primarily on synthetic channel models. While these models are valuable for initial concept validation, but they often fail to capture the complexities, imperfections, and noise characteristics of real-world V2I communication scenarios. To highlight this difference and provide critical context, Table 1 summarizes and compares previous research works based on their methodologies and dataset types.

This research stands out by operating real-world vehicular communication data rather than conventional synthetic data, concentrating on real-world challenges. The empirical dataset enables a more precise calculation of ANN models for beam prediction. Through evaluations of various ANN architectures, the research identifies the most effective one and highlights the advantages of engaging data-driven, location-aware beamforming to enhance the performance and consistency of 5G NR V2I networks.

Table 1. Summary and comparison of previous research work.

Reference	Problem Addressed	Methodology	Dataset Type	Key Findings	Limitations
Morais et al. [1]	GPS-aided beam prediction in real world	ML regression models	Real-world drive tests	GPS improves beam selection	Sensitive to GPS errors
Hassanein et al. [2]	Deep learning for 5G/6G beam prediction	DL classification	Synthetic	High accuracy in stable channels	Not validated in mobility
Minchula et al. [3]	MIMO under ICSIT/ICSIR	Analytical modeling	Synthetic	Evaluates multi-antenna behavior	Pre-5G, no beamforming
Minchula and Rao [4]	SVD-based IWFA	Optimization	-	Fast convergence	No AI or mobility
Khan et al. [5]	Precoding for massive MIMO	ML-optimized precoder	Simulator	Improved energy efficiency	Not V2X-specific
Minchula and Sasibhushana Rao[6]	SAC-channel impact	Analytical	-	Capacity improvements	No DL techniques
Ktari and Rekaya [7]	Multi-user beam mgmt	ANN	Simulator	Throughput gains	No vehicular mobility
Bilaminu et al. [8]	Dataset/codebook size impact	ML classifier	Synthetic V2I	Larger codebook hurts accuracy	No multimodal sensing
Dutta et al. [9]	Smart city MU-MIMO	MU-MIMO analysis	Realistic	Energy efficiency	Not beam prediction
Deng et al. [10]	Sub-6 GHz aided mmWave prediction	SPBPNet (DL)	Hybrid	Reduced pilot overhead	Complex
Dissanayake and Temdee [11]	CNN/ANN for 6G	CNN + ANN	Simulator	Better millimeter-wave accuracy	No mobility
Vankayala et al. [12]	Future blockage + beam prediction	CNN + LSTM	Synthetic	Good predictive accuracy	No real data
Varma et al. [13]	NN-based beamforming	ANN	Simulator	Improved BF accuracy	Not V2X-oriented
Kodam et al. [14]	Relay-based V2V under NLOS	Empirical modeling	Simulator	Better reliability	Not beam prediction

Table 1 (continue). Summary and comparison of previous research work.

Reference	Problem Addressed	Methodology	Dataset Type	Key Findings	Limitations
Jalali et al. [15]	Fast best beam selection	DL	Simulator	Lower overhead	Not V2X-oriented
Alrabeiah and Alkhateeb [16]	Vision-aided mmWave tracking	YOLO + DL block	Camera + RF	Multimodal sensing boosts accuracy	High hardware complexity
Charan et al. [17]	Sub-6 GHz aided blockage/beam prediction	DL regression	Real + synthetic	Strong cross-band correlation	Needs calibration
Proposed Work	V2I beam prediction for 5G NR mmWave bands.	ANN Comparative Analysis	Empirical	Identifies optimal architecture (LSTM)	-

The primary contributions of this research are its empirical validation utilizing a large-scale real-world V2I empirical dataset [1] for beam prediction, contrasting sharply with the predominantly synthetic or simulated data used in prior literature. Thereby providing a more realistic performance evaluation for 5G NR deployment, which inherently includes real channel impairments and noise.

Secondly, a rigorous comparative analysis of three distinct and prominent ANN architectures - CNN, RNN, and LSTM was used to determine the optimal model for position-aided beam selection in high-mobility V2I channels. Thirdly, the performance evaluation is carried out using Top-K accuracy through a quantitative metric, Power Loss (dB), which provides a measure of the effective signal strength penalty incurred by incorrect beam prediction, offering a crucial metric for real-world link reliability assessment.

Finally, in comparison with different ANN models, the LSTM architecture consistently achieves superior performance across nine distinct real-world scenarios, identifying it as the most effective deep learning solution for robust, position-aided beam alignment in dynamic 5G V2I environments. The following section outlines the proposed methodology and details the dataset utilized for beam prediction.

The third section presents the system architecture, including the neural network design and the hyperparameters employed during training. The fourth section provides a comprehensive analysis of the experimental results, including evaluation metrics and performance comparisons. Finally, the concluding section summarizes the key findings and discusses potential directions for future research.

2. Method

Predicting the optimal beam based on positional data can be effectively addressed using deep learning techniques, particularly those suited for capturing spatial and temporal patterns. This research investigates the problem using three distinct neural network architectures: CNN, RNN, and LSTM networks. A prediction function p_{Θ} , where Θ represents trainable parameters is defined. The function is trained on a dataset consisting of input-output pairs, where each input is a user's position and

the output is the corresponding optimal beam. Once trained, this function can estimate the appropriate beam for any new location input. “The parameters Θ are learned from a dataset $D = \{(g_k, p_k^*): k = 1, \dots, K\}$ which is composed of K labeled training samples. Each sample consists of the input location g_k and its ground truth optimal beamforming vector $p_k^* \in F$. Therefore, for a given position g in Eq. (1).

$$\hat{p} = p \ominus (g) \quad (1)$$

The CNN is the first model to consider. 1D CNNs are well-suited to structured sequential data, such as an ordered feature vector representing spatial or angular measurements. Their ability to detect local patterns along one dimension allows them to learn spatially consistent beamforming decisions that generalize across similar conditions. Next, RNN is employed to model sequential position data. This is useful when user mobility over time influences beam selection. RNNs process inputs in order and maintain internal memory, allowing them to make context-aware predictions.

Finally, LSTM network, an enhanced type of RNN capable of preserving information over longer sequences, is considered. LSTMs are especially effective in scenarios where the user’s movement history impacts the current beam choices, as they can capture long-term dependencies in the input data. All three models generate a set of probabilities over possible beam choices. The models are trained using the Adam optimizer and the Categorical Cross-Entropy loss function, suitable for the multi-class beam classification task. The beam with the highest predicted probability is chosen as the output, allowing the model to not only make a decision but also reflects the confidence behind it.

$$\hat{p} = p_{\hat{m}}, \hat{m} = \underset{m \in \{1, \dots, M\}}{\operatorname{argmax}} p_m \quad (2)$$

The following section explains how each model processes the positional data to compute the probability distribution and the flow diagram about different ANN models for position aided beam prediction is depicted in Fig. 1.

2.1. Convolutional neural networks

The process of using a CNN for beam prediction starts with feeding a 1D feature vector into the network, where each element holds features like signal intensity, angle information, or channel measurements arranged in an ordered manner. This input is processed through multiple 1D convolutional layers that apply filters along the sequence to detect local dependencies, with activation functions such as ReLU introducing non-linearity.

Pooling layers may follow to down sample the feature sequence while preserving key information. As the data moves deeper through the network, more complex and abstract features are learned, which are crucial for accurate beam prediction. The final layers—either convolutional or fully connected—output a flattened feature vector that is passed through a softmax layer to generate a probability distribution over possible beam directions, identifying the most suitable configuration. See Algorithm 1 for description.

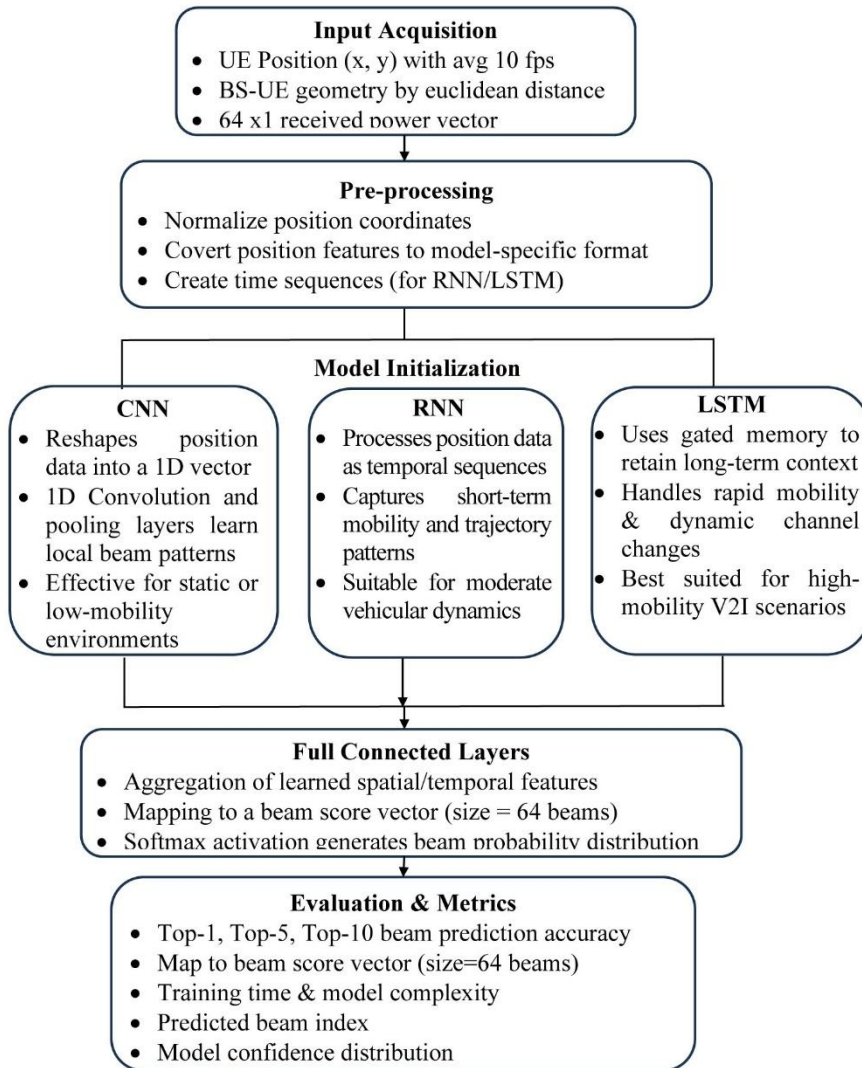


Fig. 1. Flowchart of position-aided beam prediction using ANNs in 5G V2I.

2.2. Recurrent neural networks

In an RNN-based beam prediction flow, each input position-described by features like location, signal intensity, or direction-is input to the network one at a time. As the RNN processes each position, it updates its hidden state to retain relevant information from earlier inputs, effectively building a contextual understanding of the sequence. This evolving hidden state helps the network capture relationships between positions and supports accurate prediction of the beam configuration for each step. At every position, the network generates a prediction using the current hidden state, and a softmax layer is applied to output a probability distribution over the possible beam directions, enabling context-aware and sequentially informed beam selection. See Algorithm 1 for description.

Algorithm 1

• **Input:** Position (x_1, x_2) , Beam Index (y)

• **Normalization:** $x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)}$

Training of ANN:

Convolutional Neural Network

Require: g , trained CNN model F_θ

$\hat{P} \leftarrow F_\theta(g)$ ▶ Predict beam probabilities

$\hat{m} \leftarrow \operatorname{argmax}_m \hat{P}_m$ ▶ Select beam with highest probability

Return \hat{P}, \hat{m} ▶ Return prediction

Recurrent Neural Network

Require: $G = [g_1, g_2, \dots, g_T]$, trained RNN model R_θ

$h_0 \leftarrow 0$ ▶ Initialize hidden state

for $t \leftarrow 1$ to T do

$h_t \leftarrow R_\theta(g_t, h_{t-1})$ ▶ Update hidden state with input g_t

end for

$\hat{P} \leftarrow \operatorname{Softmax}(Wh_T + b)$ ▶ Predict beam probabilities
from hidden state

$\hat{m} \leftarrow \operatorname{argmax}_m \hat{P}_m$ ▶ Select beam with highest probability

Return \hat{P}, \hat{m} ▶ Return prediction

Long Short-Term Memory

Require: $G = [g_1, g_2, \dots, g_T]$, trained LSTM model L_θ

$h_0 \leftarrow 0, c_0 \leftarrow 0$ ▶ Initialize hidden and cell states

for $t \leftarrow 1$ to T do

$(h_t, c_t) \leftarrow L_\theta(g_t, h_{t-1}, c_{t-1})$ ▶ Update LSTM states with input g_t

end for

$\hat{P} \leftarrow \operatorname{Softmax}(Wh_T + b)$ ▶ Predict beam probabilities
from final hidden state

$\hat{m} \leftarrow \operatorname{argmax}_m \hat{P}_m$ ▶ Select beam with highest probability

Return \hat{P}, \hat{m} ▶ Return prediction

2.3. Long short-term memory

In beam prediction using an LSTM network, each input position-described by features such as location, beam intensity, or direction-is processed in sequence. The LSTM uses its internal memory and gating mechanisms (input, forget, and output gates) to retain important contextual information while filtering out irrelevant data. This sequential processing enables the model to learn spatial or temporal dependencies across the input positions. As it moves through the sequence, the LSTM updates its hidden state to accumulate relevant information, which is then used to predict the optimal beam configuration at each step.

These predictions are typically generated by passing the hidden state through a dense layer with softmax activation, resulting in a probability distribution over possible beam directions and allowing for accurate, context-driven beam selection. See Algorithm 1 for description. Top-k accuracy measures the proportion of test samples for which the correct label appears among the Top-k predicted outputs. Top-1 accuracy reflects the strictest case, where only the most confident prediction is considered, indicating exact matches.

Top-2 accuracy includes the Top two predictions, allowing more flexibility by considering the second-best guess as well. This continues through Top-3, Top-4, and

Top-5, each progressively expanding the set of acceptable predictions. As k increases, the accuracy typically improves, offering a more tolerant evaluation metric, especially useful when exact prediction is challenging but approximate correctness is still valuable. This metric is widely used in classification tasks like beam prediction to assess how often the true value is captured within the Top- k guesses.

3. Architecture and Parameters

This work utilizes the DeepSense dataset [1], a large-scale, multi-modal real-world dataset developed by the Wireless Intelligence Lab, which integrates synchronized GPS coordinates, camera images, including received signal power. Scenario sample ranges from 854 to 5964 samples, including daytime, nighttime, and rainy conditions (e.g., Scenarios 1: 2411, 2: 2974, 3: 1487, 4: 1867, 5: 2300, 6: 915, 7: 854, 8: 4043, 9: 5964) with capture rates of average 5.5 to 7.7 FPS. These diverse environments support beam prediction, positioning, and multi-modal learning benchmarks under various real urban conditions.

Specifically, calibrated GPS positions and beam power data from Scenarios 1 to 9 are used. The data was collected across various locations on the ASU campus under different environmental conditions. A moving vehicle equipped with a 60 GHz omnidirectional mmWave transmitter and GPS served as the UE, a 16-element phased array captures incoming signals through an over-sampled codebook comprising 64 predefined beams. The experiments were implemented using Python with the TensorFlow/Keras deep learning framework.

At a sampling rate of approximately 10 Hz, the BS swept through a predefined codebook, recording received power across all beams. The dataset captures the vehicle passing the BS from multiple directions, and position-beam associations are visualized as scatter plots, where each GPS point is color-coded according to the index of the strongest received beam out of 64 possible azimuthal beams. The dataset for each scenario was partitioned into a stratified split of 60% for training, 20% for validation, and 20% for testing to support model development and evaluation. This stratified split, with equal allocation for validation and test sets, follows standard machine learning conventions.

Preprocessing of dataset

The architecture of the 1D CNN exhibited in this research is structured to effectively perform beam prediction by leveraging layered feature extraction and classification stages. The parameters are presented in Table 2. It begins with three 1D convolutional layers that progressively increase the number of filters—from 32 to 128—enabling the structure to capture local sequential or angular feature patterns at numerous levels of abstraction. Each convolutional layer is followed by a ReLU activation and a 1D max pooling operation, which reduces the length of the feature sequence progressively compressing the input representation.

The resulting feature map, containing 128 channels, is then flattened and passed through two fully connected layers: the first maps the learned feature representation to 512 neurons, whereas the second maps it to the final output dimension for beam prediction. To improve generalization and prevent overfitting, a dropout layer with a probability of 0.5 is applied after the first fully connected layer. This 1D CNN

architecture is built to efficiently capture sequential spatial or angular dependencies and provide accurate predictions in mmWave beamforming tasks.

Table 2. Different parameters of ANN.

Parameters	CNN	RNN	LSTM
LEARNING RATE	0.01	0.01	0.01
BATCH SIZE	32	32	32
NUMBER OF EPOCHS	60	60	60
L2 REGULARIZATION	$1e^{-5}$	$1e^{-5}$	$1e^{-5}$
HIDDEN SIZE	-	128	128
NUMBER OF LAYERS	-	2	2

The two models (RNN, LSTM) are created for Top-K beam prediction based on two-dimensional positional inputs. The training parameters are described in Table 2. The hyperparameters, such as the Learning Rate (0.01) and Batch Size (32), were chosen as standard starting points refined through empirical testing to ensure convergence without premature stagnation. Each model reformats the input data to encompass a temporal measurement, permitting sequence modelling. The primary model leverages a RNN architecture with customizable hidden units and multiple layers to capture transient designs. The hidden size of 128 and 2 layers were selected to balance complexity and the need to model vehicular trajectory over time.

L2 regularization is applied to both RNN and LSTM to further mitigate overfitting. After the sequence is processed, the output is passed over an actuation layer, taken after by dropout to minimize overfitting, and after that through a fully associated layer to generate the ending expectation. It uses a reset hidden state at the start of each sequence. In contrast, the second model employs a LSTM network, which also permits customization of hidden size and layer depth. This model initializes both the hidden and cell states and follows a similar processing pipeline, finishing with a linear layer to deliver the yield. Whereas both models can learn from sequential information, the LSTM's design allows it to hold longer-term circumstances, making it specifically convincing for modelling complicated, time-varying situations.

4. Results and Discussion

This section evaluates the performance of various algorithms using the real-world dataset. When using DL approach to select the best beam, it must ensure that not only to check the first guess but also it must consider how the ANN models will choose the best beam from the Top 3 or Top 5 guesses. The Fig. 2 illustrates the Top-k accuracy comparison between the three models for Scenario 2. The results clearly show that LSTM and RNN, which are sequence-based models, significantly outperform the CNN across all k values.

Specifically, LSTM achieves the highest Top-5 accuracy, outperforming RNN by 4.24% and CNN by a substantial 45.51%, highlighting the critical role of temporal context. LSTM demonstrates superior generalization across varying vehicular scenarios and its ability to capture long-term dependencies within the sequential position data. RNN exhibits a strong performance, underscoring the benefit of temporal modelling, while 1D CNN, which is limited to capturing features from local sequential patterns, demonstrates the lowest accuracy.

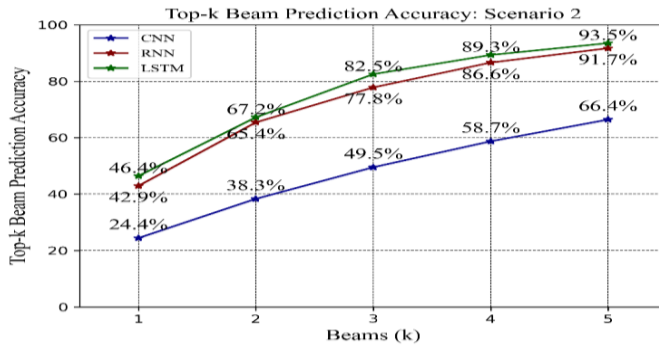


Fig. 2. Top-k accuracy comparison between algorithms for Scenario 2.

Figure 3 presents a comparison of Top-2 beam prediction accuracy across the nine distinct real-world V2I scenarios. This analysis serves as an architectural ablation study and a critical reliability check, validating the models' generalization capability across varied environmental conditions. It clearly demonstrates that LSTM achieves the highest performance in most scenarios. It gives the highest accuracy in 7 out of 9 cases. RNN comes close behind LSTM and often performs better than CNN. CNN performs the worst in almost all scenarios, with very low accuracy in some (e.g., only 8.5% in Scenario 1 and 8.0% in Scenario 5).

The consistently high accuracy of LSTM and RNN across all nine replications (scenarios) indicates their robustness and reliable generalization, confirming that sequence-based deep learning models are necessary for high-mobility environments. Unlike CNN, the temporal models demonstrate superior consistency and do not require scenario-specific tuning to generalize well across different vehicular trajectories and channel variations. Even though each model was used with its best possible settings for each scenario, LSTM and RNN still came out ahead. This proves that neural networks can better learn patterns from the data without needing special adjustments.

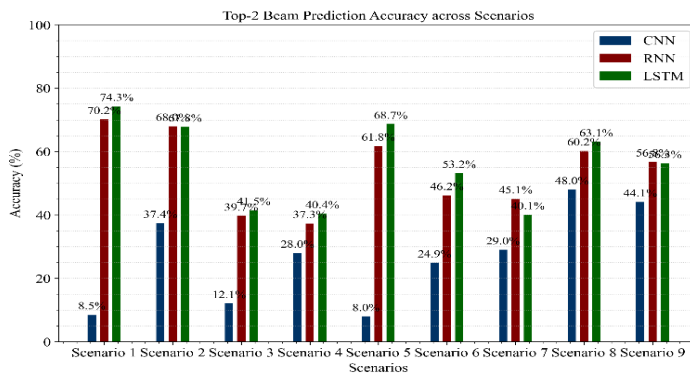


Fig. 3. Top-2 beam prediction accuracy [%] across Nine V2I Scenarios (A measure of model reliability).

The Fig. 4 bar chart illustrates a " Top-k Beam Frequency Distribution". It clearly demonstrates that some beams were favoured over others; for example, the solution

at Beam Index 42 was the most frequently chosen candidate, appearing over 120 times. Other beams, such as those near indices 8, 11, and 61, were also popular but less dominant. This visualization provides a clear snapshot of an algorithm's behaviour, highlighting which solutions it consistently identified as the most promising during its search process.

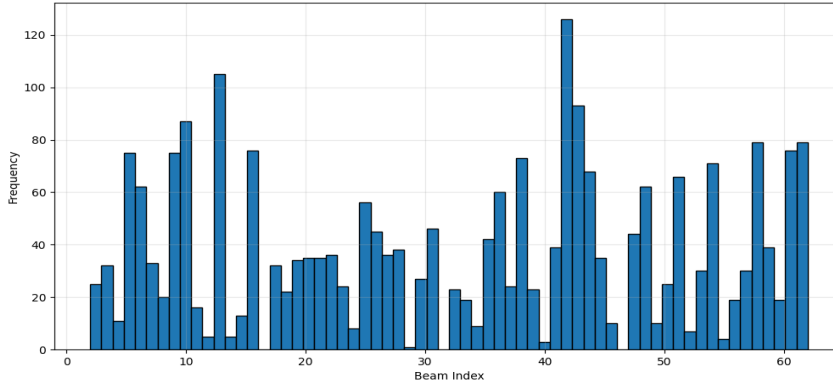


Fig. 4. Top-k beam frequency distribution.

Figure 5 presents the beam prediction maps for CNN, RNN, and LSTM models for Scenario 2. The visualization of the predicted beam indices across the geographical area provides qualitative insight into the model's learned spatial continuity. The LSTM prediction is the most spatially consistent, with smooth and organized patterns, suggesting a strong understanding of the underlying physical channel correlations and vehicular trajectory. The RNN also shows patterns, but they are more noisy and uneven. In sharp contrast, the CNN prediction appears spatially inconsistent and noisy, with no clear structure.

This qualitative evidence strongly supports the quantitative findings, demonstrating that LSTM provides the most robust and reliable beam prediction output, while CNN is unsuitable for this dynamic environment. Power loss shows how much prediction performance is lost compared to perfect accuracy. This metric serves as a critical quantitative indicator of link reliability loss. It is measured in dB. The higher the power loss, the worse the model is doing. Lower power loss means the model is making better predictions. It provides a practical measure of the signal power penalty incurred when the predicted beam is not the optimal beam, but one of the Top-K alternatives. Here, power loss is defined as:

$$L_p = \frac{1}{N} \sum_{k=1}^N -10 \cdot \log_{10} \left(\frac{A_k}{100} \right) \quad (3)$$

where L_p is overall power loss in dB, A_k is accuracy at beam k (in %) and N is the total number of Top- k values.

The Fig. 6 presents a comparison of how three neural network models-CNN, RNN, and LSTM-perform in terms of power loss (in decibels) across nine different V2I communication scenarios. The scenarios (1-9) in Table 3 represent distinct, real-world data environmental conditions, ensuring the robustness and generalization of the models. The data demonstrates that CNN consistently experiences the highest power loss in all tested scenarios, serving as a clear indicator of low link reliability

and high penalty from incorrect beam choices. Notably, in Scenario 5, CNN reaches a peak loss of around 8.8 dB, and similarly high values are seen in Scenarios 3 and 4 with approximately 6.9 dB and 6 dB, respectively. This suggests that CNN, despite its strength in capturing spatial features, is less suitable for minimizing power loss in dynamic vehicular environments.

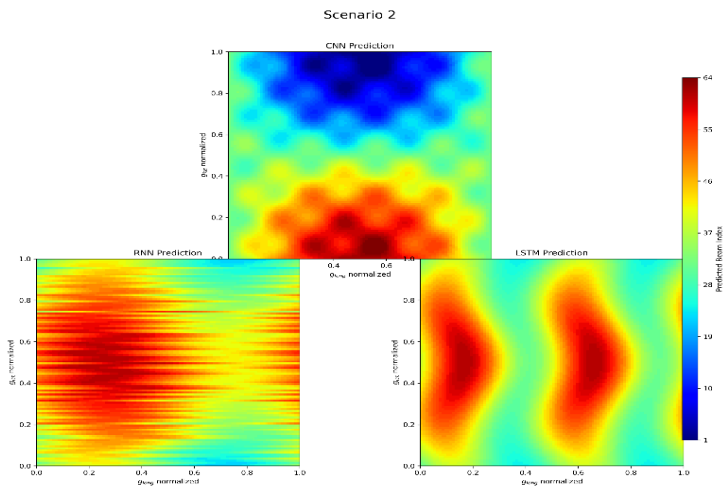


Fig. 5. Scenario 2 Prediction Maps (CNN, RNN, LSTM) Demonstrating Spatial Consistency of Predicted Beam Indices.

Table 3. Power loss across scenarios 1-9 as per ref. [1].










Scenario		CNN (dB)	RNN (dB)	LSTM (dB)
1	 V2I, SLC, Day-time Scenario	5.77	1.24	0.94
2	 V2I, SLC, Night-time Scenario	3.90	1.53	1.34
3	 V2I, SLC, Day-time Scenario	6.90	2.24	2.15
4	 V2I, SLC, Night-time Scenario	6.02	2.34	2.22
5	 V2I, SLC, Night-time Scenario, Rainy Weather Condition	8.81	1.46	1.31

Table 3 (continue). Power loss across scenarios 1-9 as per ref. [1].

Scenario		CNN (dB)	RNN (dB)	LSTM (dB)
6	 V2I, SLC, Day-time Scenario	5.58	1.55	1.33
7	 V2I, SLC, Day-time Scenario	3.74	3.21	3.10
8	 V2I, SLC, Day-time Scenario	1.66	1.86	1.66
9	 V2I, SLC, Day-time Scenario	3.90	2.15	2.08

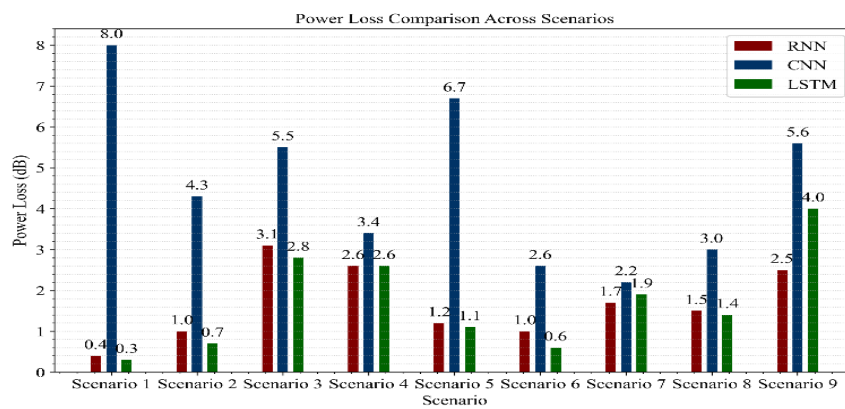


Fig. 6. Power loss comparison (in dB) across nine V2I scenarios (lower is better, indicating higher link reliability).

In contrast, RNN models demonstrate a more moderate level of power loss, with values generally lower than CNN but still higher than LSTM. In Scenarios 3 and 4, RNN records losses of about 2.3-2.4 dB, reflecting its capacity to process sequential data more effectively than CNN. However, in more variable conditions such as Scenario 7, RNN reaches about 3.2 dB, indicating some limitations under increased complexity. LSTM models show the most favourable performance, with the lowest power loss across all scenarios.

For instance, it achieves values as low as 0.9 dB in Scenario 1 and 1.3 dB in Scenario 5. Even in more complex situations, LSTM maintains its efficiency, rarely exceeding 2.1 dB. Quantitatively, LSTM reduces the average power loss by approximately 2.9 dB compared to CNN and maintains an average advantage of 0.2 dB over RNN across all scenarios. This demonstrates LSTM’s strength in managing temporal patterns and long-term dependencies, making it the most reliable model for beam prediction in high-mobility V2I networks.

The statistical results in Table 4 demonstrates that the LSTM model is superior in performance by achieving higher mean accuracies while maintaining controlled variability, reaching 38.09% Top-1 accuracy with $\sigma \approx 9.51\%$ and 82.97% Top-5 accuracy with $\sigma \approx 9.81\%$. This consistent improvement from Top-1 to Top-5, coupled with stable standard deviation across all Top-K levels, highlights the robustness and scalability of LSTM for real-world mmWave beam alignment.

In comparison, the RNN model also exhibits stable behavior, achieving 35.93% Top-1 accuracy with $\sigma = 8.01\%$ and 80.90% Top-5 accuracy with $\sigma = 9.63\%$, indicating reliable and consistent beam ranking across scenarios. However, the higher mean accuracy gains of LSTM amounting to an absolute improvement of approximately 2-4 percentage over RNN and more than 30 percentage over CNN establish it as the most effective solution across 22,814 real-world V2I samples spanning 9 scenarios, particularly in high-mobility environments.

Table 4. Mean (μ) and standard deviation (σ) of Top-K Accuracy (%).

Model	Top 1		Top 2		Top 3		Top 4		Top 5	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
LSTM	38.1	9.51	56.9	12.6	69.5	12.8	77.7	11.6	82.9	9.81
RNN	35.9	8.01	53.4	11.4	65.9	11.4	74.9	10.4	80.9	9.63
CNN	17.8	7.02	28.4	10.2	36.8	12.3	44.1	13.5	50.1	14.6
KNN	33.9	10.9	50.2	13.8	62.1	13.8	66.7	12.1	67.2	11.7
LT	31.6	8.3	48.4	11.3	59.1	13.5	65.6	14.9	69.5	14.7

5. Conclusion

This research gives a comprehensive evaluation of different deep learning models—specifically 1D CNN, RNN, and LSTM networks for Top-K beam prediction in dynamic vehicular circumstances, a straightforward chore for authorizing strong mmWave communication. The comparative examination, validated using real-world empirical V2I data, unequivocally uncovers that the LSTM architecture consistently outperforms other approaches across all nine test scenarios. LSTM achieved the best overall performance, reaching 38.04% Top-1 accuracy and 82.96% Top-5 accuracy, and demonstrated superior robustness by achieving the lowest average Power Loss (dB).

LSTM's ability to hold long-term dependencies and adapt to the sequential nature of vehicular movement makes it the most reliable and compelling model for real-world deployment, surpassing the RNN's performance by 4.24% and the CNN's performance by 45.51% in Top-5 accuracy. RNNs also present a favourable trade-off between computational efficiency and prediction accuracy. In contrast, CNNs, although capable of depicting local sequential features, exhibit limited performance in highly dynamic environments, time-varying settings due to their constrained capacity for modelling temporal properties.

Despite demonstrating the effectiveness of the LSTM model, this work is subject to certain limitations. First, the analysis is constrained by the scope of the DeepSense dataset, which covers a specific urban-campus environment (ASU) and may not fully generalize to radically different geographies or channel conditions (e.g., dense urban canyons, rural or highway scenarios). Second, while the use of empirical data inherently includes real channel impairments and noise, the

evaluation does not explicitly incorporate external noise models or perform formal hypothesis testing on the results.

Finally, this research focuses solely on beam prediction accuracy and Power Loss, omitting an analysis of the computational latency, which is a critical factor for real-time URLLC applications. The current framework can be further enhanced by incorporating controlled noise models, statistical significance analysis, and robustness evaluation under synthetic perturbations to further strengthen the analytical rigor. By enabling autonomous, data-driven beam prediction using real-world vehicular context, the proposed framework supports the objectives of the ITU-T Autonomous Massive MIMO use case (FG-AN-usecase-008) for 5G and beyond.

Future research should focus on three specific, actionable directions. First, investigating hybrid architectures (e.g., CNN–LSTM) to simultaneously exploit local feature patterns from ordered beam-related measurements (1D CNN) and temporal dependencies from sequential position history (LSTM) may yield further performance gains. Second, research should explore multimodal data fusion by incorporating additional sensor data available in the V2I environment, such as camera images or LiDAR data, to better mitigate blockage events. Third, a crucial step for real-world deployment is conducting a thorough analysis of the computational complexity and latency of the LSTM model to ensure it meets the strict real-time processing constraints required for 5G NR URLLC vehicular communications.

Acknowledgments

Nomenclatures

A_k	Top-k accuracy
b	Bias vectors in the neural network layers
c_t	Cell state vector at time step t (LSTM)
Fm	Learned Feature Map Vector
g_k	Input Location
N	Total number of Top beams considered
P	Total Power Loss
p^*	Optimal Beamforming Vector
W	Weight matrices in the neural network
y	Beam Index
θ	Trainable parameters (weights and biases) of the ANN

Abbreviations

ASU	Arizona State University
ANN	Artificial Neural Networks
BS	BaseStation
CNN	Convolutional Neural Network
CSI	Channel State Information
dB	decibels
DL	Deep Learning
eMBB	Enhanced Mobile Broadband
GPS	Global Positioning System
LSTM	Long Short-Term Memory

mmWave	millimeter-wave
MIMO	Multiple-Input and Multiple-Output
NR	New Radio
RNN	Recurrent Neural Networks
SLC	Street-Level Communication
GPS	Global Positioning System
UE	User Equipment
URLLC	Ultra-Reliable Low-Latency Communication
V2I	Vehicle to Infrastructure

References

1. Morais, J.; Bchboodi, A.; Pezeshki, H.; and Alkhateeb, A. (2023). Position-aided beam prediction in the real world: How useful GPS locations actually are? *Proceedings of the ICC 2023 IEEE International Conference on Communications*, Rome, Italy, 1824-1829.
2. Hassanein, S.; Shouman, O.; Yaacoub, E.; Al-Ali, A.; and Khattab, T. (2024). Beam prediction in 5G and 6G systems: A deep learning approach. *Proceedings of the 2024 6th International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, Istanbul, Turkiye, 1-6.
3. Minchula, V.K.; Rao, G.S.; and Vijay, C. (2018). *Multi-antenna system performance under ICSIT and ICSIR channel conditions*. In Bhateja, V.; Coello Coello, C.A.; Satapathy, S.C.; and Pattnaik, P.K. (Eds.), *Intelligent Engineering Informatics*. Springer Singapore, 505-511.
4. Minchula, V.K.; and Rao, G.S. (2018). SVD-based IWFA for next generation wireless MIMO communication. *ICT Express*, 4(3), 171-174.
5. Khan, A.A.; Almuzaini, K.K.; Macedo, V.D.J.; Ojo, S.; Minchula, V.K.; and Roy, V. (2023). MaReSPS for energy efficient spectral precoding technique in large scale MIMO-OFDM. *Physical Communication*, 58, 102057.
6. Minchula, V.K.; and Sasibhushana Rao, G. (2019). *SAC channel effects on MIMO wireless system capacity*. In Panda, G.; Satapathy, S.C.; Biswal, B.; and Bansal, R. (Eds.), *Microelectronics, electromagnetics and telecommunications*. Springer Singapore, 759-765.
7. Ktari, A.; and Rekaya, G. (2024). Neural networks for multi user beam management in mmWave massive MIMO. *Proceedings of the 2024 IEEE Middle East Conference on Communications and Networking (MECOM)*, Abu Dhabi, United Arab Emirates, 94-99.
8. Biliaminu, K.K.; Busari, S.A.; Bastos, J.; and Rodriguez, J. (2024). Impacts of dataset and codebook sizes on ML-Driven beam prediction for mmWave V2I communication. *Proceedings of the 2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Leeds, United Kingdom, 1-6.
9. Dutta, P.; Ramyasree, J.; Sridhar, V.; Minchula, V.K.; Mohanta, H.C.; Mahfoudh, S.; Shah, S.B.H.; and Singh, S.P. (2022). Evaluating the efficiency of non-orthogonal MU-MIMO methods in smart cities technologies & 5G communication. *Sustainability*, 15(1), 236.

10. Deng, W.; Li, M.; Liu, Y.; Zhao, M.-M.; and Lei, M. (2024). Enhancing mmWave beam prediction through deep learning with sub-6 GHz channel estimate. *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, Dubai, United Arab Emirates, 1-6.
11. Dissanayake, D.M.C.; and Temdee, P. (2023). Towards 6G: Beam prediction using convolutional neural network and artificial neural network. *Proceedings of the 2023 7th International Conference on Information Technology (InCIT)*, Chiang Rai, Thailand, 392-396.
12. Vankayala, S.K.; Gollapudi, S.K.S.; Jain, B.; Yoon, S.; Mihir, K.; Kumar, S.; Sudheer Kumar, H.U.; and Kommineni, I. (2023). Efficient deep-learning models for future blockage and beam prediction for mmWave systems. *Proceedings of the NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*, Miami, FL, USA, 1-8.
13. Varma, P.; Patil, S.; and Gupta, H. (2024). Optimizing millimeter-wave beamforming in 5G and beyond with neural networks. *Proceedings of the 2024 1st International Conference on Sustainable Computing and Integrated Communication in Changing Landscape of AI (ICSCAI)*, Greater Noida, India, 1-11.
14. Kodam, S.; Minchula, D.V.K.; Reddy, A.S.; Chalamalasetty, K.; and Tammireddi, R.S. (2024). Optimizing relay based V2V communication in Non-Line-of-Sight scenarios: Pathloss modelling and strategies. *Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, 1-7.
15. Jalali, J.; Roa, J.; Song, Y.; Zhao, R.; and Sheen, B. (2024). Fast best beam prediction and overhead reduction for 6G networks: A deep learning approach. *Proceedings of the 2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)*, Singapore, 01-06.
16. Alrabeiah, M.; and Alkhateeb, A. (2020). Deep learning for mmWave beam and blockage prediction using sub-6 GHz channels. *IEEE Transactions on Communications*, 68(9), 5504-5518.
17. Charan, G.; Alrabeiah, M.; and Alkhateeb, A. (2021). Vision-aided 6G wireless communications: Blockage prediction and proactive handoff. *IEEE Transactions on Vehicular Technology*, 70(10), 10193-10208.