IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 October 2025, accepted 18 October 2025, date of publication 30 October 2025, date of current version 5 November 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3626953

RESEARCH ARTICLE

GPS-Based Beam Prediction Using Lightweight Deep Learning Models for mmWave Networks

MUHAMMAD HASEEB HASHIR[1], MEMOONA[1], AND SUNG WON KIM[2]
[1]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea
[2]School of Computer Science and Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea

Corresponding author: Sung Won Kim (swon@yu.ac.kr)

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2021R1A6A1A03039493, and in part by the NRF grant funded by Korea Government [Ministry of Science and ICT (MSIT)] under Grant NRF-2022R1A2C1004401.

ABSTRACT Millimeter-wave (mmWave) communication systems use narrow, directional beams, but exhaustive beam training is costly in dynamic settings. Prior location-assisted methods often rely on synthetic data that ignores the noise of real-world GPS measurements. We propose a unified framework that first denoises GPS trajectories with Gaussian-process regression and then predicts beams using a bidirectional long short-term memory network with an attention mechanism. Across nine real-world scenarios, our approach improves Top-k accuracy by up to 36%, reduces received-power loss by more than 1 dB, and cuts beam-training overhead by up to 95%. These results highlight the effectiveness of the proposed framework in bridging the gap between simulation-driven research and real-world mmWave beam alignment.

INDEX TERMS Beam prediction, millimeter-wave, deep learning.

I. INTRODUCTION

Millimeter frequencies have emerged as a fundamental component of fifth-generation (5G) cellular networks, playing a pivotal role in meeting the demanding requirements for high data throughput and extremely low latency [1]. Recent advances in 6G wireless communications have further highlighted the critical importance of mmWave systems, with emerging applications in cellular-connected autonomous aerial vehicles (AAV) and vehicle-to-infrastructure (V2I) communications requiring even more sophisticated beam management strategies [2], [3]. A key characteristic of mmWave communication is the use of narrow, highly directional beams—commonly referred to as beamforming (BF)—which are essential for mitigating the substantial isotropic path loss inherent to these frequency bands [4], [5].

Conventional beam alignment techniques that utilize exhaustive search algorithms offer high accuracy but are generally unsuitable for real-time deployment due to their substantial time and resource demands. The process of beam alignment introduces significant training overhead, particularly in dynamic scenarios where frequent changes in the relative positions of the Base Station (BS) and User Equipment (UE) occur due to high mobility. Contemporary research has demonstrated that artificial intelligence and machine learning approaches are becoming indispensable for addressing the increasing complexity of beam management in modern wireless systems, particularly as antenna arrays become larger and operating frequencies move toward sub-THz bands [6]. This challenge is amplified in mission-critical applications and highly mobile scenarios such as drone communications, where continuous service provision requires innovative beam prediction architectures [7], [8]. Recent advances in attention mechanisms and transformer architectures have shown promise in addressing these challenges, with attention-enhanced learning models achieving Top-5 beam prediction accuracies exceeding 90% across multiple future time slots [9], [10].

Many previous studies have leveraged user positioning information to support beam alignment in mmWave systems [11], [12], [13], [14]. Leveraging user location data, such as that obtained via GPS, presents a promising opportunity to

The associate editor coordinating the review of this manuscript and approving it for publication was Adao Silva.

expedite the beam alignment process and significantly reduce associated overhead. However, most earlier works relied on synthetic datasets, typically under idealized assumptions that fail to reflect the complexity and noise present in real-world scenarios. Synthetic datasets often overlook the intricate dynamics of real-world environments, resulting in models that perform well in simulations but degrade in deployment. Moreover, mmWave signals are highly vulnerable to obstructions caused by human bodies, hand movements, and construction materials [15], [16], which are inadequately represented in synthetic studies. While machine learning has been explored as a potential solution, existing approaches inadequately address the inherent noise in real GPS data, limiting their practical applicability. This highlights a notable research gap in developing robust deep-learning models that utilize real positional information with effective denoising for precise beam selection.

Method Summary: To bridge these gaps, we propose a pipeline that first applies Gaussian Process (GP) regression to denoise noisy GPS trajectories, then leverages sequential modeling with Bidirectional Long Short-Term Memory (Bi-LSTM) and attention mechanisms to capture temporal dependencies, and finally employs a dual-head architecture to simultaneously classify candidate beams and regress link quality. At inference, the model shortlists a small set of beams based on predicted probabilities and selects the optimal one using the quality estimates. This unified pipeline enables robust, lightweight, and real-world–validated beam prediction without reliance on synthetic datasets.

Main contributions:

- **Unified GPS-Enhanced Framework:** We introduce a unified framework that combines Gaussian Process regression with recurrent neural architectures (Bi-LSTM and Bi-LSTM + Attention), enabling robust denoising of GPS trajectories and accurate beam prediction in a single pipeline.
- **Advanced GPS Denoising:** We employ GP regression to smooth noisy GPS data, thereby enhancing the stability and reliability of location-based features used for beam prediction—addressing a critical gap in existing location-based approaches.
- **Lightweight Architecture Design:** We design two lightweight recurrent neural architectures—Bi-LSTM and Bi-LSTM with Attention—that jointly perform candidate beam classification and link-quality regression using sequential GPS inputs, overcoming computational limitations of existing methods.
- **Real-World Validation:** We develop a real-time beam prediction framework that operates solely on consumer-grade GPS data and is trained entirely on real-world measurements, eliminating reliance on synthetic simulations that plague existing approaches.
- **Comprehensive Performance Analysis:** We conduct extensive validation across nine diverse DeepSense6G scenarios, demonstrating substantial improvements in Top-k accuracy, power efficiency, and beam-training overhead reduction compared to baseline approaches.

II. RELATED WORK

In recent years, the issue of achieving efficient beam alignment in millimeter-wave (mmWave) communication systems has emerged as a prominent area of research focus. Traditional beamforming techniques frequently depend on exhaustive beam search procedures, which, although capable of identifying optimal beam directions, impose considerable computational and time overhead. This makes them ill-suited for dynamic or real-time applications, especially in high-mobility scenarios. Their reliance on scanning large beam codebooks for every transmission further exacerbates latency and energy consumption. These constraints highlight the urgent need for more intelligent and adaptive beam alignment frameworks—preferably those that leverage contextual data, such as user location or environmental cues, to significantly reduce overhead while maintaining alignment accuracy. With this perspective, recent research efforts have explored the idea of utilizing user location data to predict suitable beam directions, aiming to significantly reduce the beam search space and associated overhead. In [17], the authors address the challenge of achieving efficient beam alignment in mmWave systems, where traditional exhaustive search techniques are often impractical due to high overhead. To mitigate this, they propose a position-assisted beam alignment strategy tailored for the 28 GHz band, validated through ray tracing simulations. A major contribution of their work is the development of an antenna alignment method that incorporates anticipated user location uncertainty, with reference to database resolutions of 10 meters and 4 meters, respectively. Simulation findings reveal that reference point spacing can be extended up to 2 meters with minimal impact on performance, underscoring the resilience of the proposed approach under moderate positional uncertainty. In [18] the authors tackle the challenge of high beam training overhead in millimeter-wave (mmWave) and terahertz (THz) communication systems, particularly in highly mobile scenarios such as drone communications. To address this, they propose a machine learning-based framework that integrates auxiliary sensory inputs—specifically visual and positional data—to enable fast and accurate beam prediction. Reference [12] addresses the limitations of location-aided beam alignment in scenarios where the receiver's orientation is not fixed, such as pedestrian applications. The paper proposes a beam selection framework using deep neural networks that integrate receiver location and orientation information for improved prediction accuracy. The model generates a prioritized set of candidate beam pairs, effectively minimizing the overhead associated with beam alignment. In this study [19] the authors propose a position-based beam prediction and selection strategy tailored for mmWave cellular systems, particularly targeting high-mobility users. Their method utilizes user positional data—specifically distance and angles of arrival and departure—to predict and

select the next optimal beam. This study [20] addresses the challenge of efficient beam alignment in dynamic V2V mmWave communication by introducing a transformer-based framework that leverages GPS and camera data for proactive, low-overhead beam tracking. This work [21] proposes a GPS-based beam selection method using a hybrid NN-KNN model to improve downlink beam alignment in mmWave networks, offering a low-overhead alternative to traditional exhaustive search techniques as networks advance toward 6G.

In [22] the authors propose an energy-efficient federated learning (EFML) framework for mmWave vehicular networks, achieving energy savings while meeting data rate demands. They also introduce a MAC-layer beam alignment prediction algorithm to reduce beamforming overhead.

In [23], the authors develop an optimized D-LSTM model combining DNN and LSTM to enhance channel estimation in mmWave MIMO systems, achieving improved NMSE and spectral efficiency over traditional methods.

In the recent study [24] the authors explore the application of deep learning techniques for predicting optimal mmWave beams to improve communication efficiency within heterogeneous networks. They introduce a dual-band fusion approach that leverages both sub-6 GHz channel information and mmWave signal measurements to facilitate effective millimeter-wave base station selection and beam training. This strategy significantly reduces training overhead while enhancing the overall performance of the network.

This work [25] proposes a closed-loop low-complexity approach for mmWave communication in smart industrial platforms, focusing on throughput and reliability. Using beam pool-based coordination and multi-agent reinforcement learning, the method enables distributed user association and soft switching, reducing complexity while enhancing system performance.

Recent works have addressed beam management challenges in next-generation wireless systems. One study examined initial access, beam tracking, and failure recovery in 5G and beyond, while also considering architectural aspects, AI-driven solutions, cooperative strategies, and standardization efforts [26]. Another focused on mmWave and THz communications for 6G, analyzing the role of AI, RIS, and ISAC [27]. AI-based methods were classified into supervised, reinforcement, and collaborative approaches, with RIS-assisted and sensing-driven predictive strategies identified as promising for high-mobility scenarios. Distinctive contributions include extending analysis to THz frequencies, exploring multi-agent learning, and unifying AI–RIS–ISAC within a single framework. Together, these works highlight open challenges in achieving lightweight and adaptive beam prediction for future networks.

Large language models have recently been explored for mmWave beam prediction. One study reformulated the task as time-series forecasting by reprogramming historical beam indices and AoDs into text-like sequences [28]. A cross-variable attention module and prompt-as-prefix mechanism were employed to enhance contextual learning. The approach achieved superior accuracy and robustness across antenna configurations, base stations, and frequencies, showing the potential of foundation models to generalize beam prediction beyond task-specific architectures. Building on this direction, another work proposed M2BeamLLM, a multi-modal framework that integrates images, radar, LiDAR, and GPS for V2X communication scenarios [29]. By combining multi-modal feature alignment with GPT-2 fine-tuning, the model significantly outperformed deep learning baselines in both standard and few-shot tasks, achieving over 13% gains in Top-1 accuracy. Ablation studies further revealed that prediction accuracy improves with modality diversity and deeper LLM fine-tuning, while inference remained efficient enough for real-time deployment. These findings underscore the promise of LLM-based multi-modal strategies for scalable and reliable beam prediction in dynamic environments.

In addition to supervised deep learning techniques, reinforcement learning (RL) has been extensively investigated for tackling control and coordination problems in dynamic environments. For instance, Li et al. [30] proposed an arbitration RL framework for heterogeneous multi-agent systems, where improved Q-functions dynamically balance on-policy and off-policy learning, achieving enhanced data efficiency and robust synchronization without requiring observers. Similarly, Wang et al. [31] introduced a two-dimensional (2D) off-policy model-free Q-learning algorithm for batch processes with actuator failures, which employs state increments in the batch direction and output errors in the time direction as novel state variables to achieve output-feedback fault-tolerant control.

More specifically, in the context of mmWave beam alignment, several recent studies have explored RL-based approaches but revealed significant limitations. Tandler et al. investigated the applicability of deep reinforcement learning algorithms to the adaptive initial access beam alignment problem for mmWave communications using proximal policy optimization, but demonstrated that the chosen off-the-shelf deep reinforcement learning agent fails to perform well when trained on realistic problem sizes. [32]. Van Huynh et al. developed a lightweight parallel reinforcement learning approach for optimal beam association in high mobility mmWave vehicular networks, specifically designed to address the computational overhead limitations of traditional RL methods [33].

Furthermore, a comprehensive literature survey by Brilhante et al. highlighted that while RL comprises an agent interacting with an environment and receiving positive or negative reinforcement responses through iterative exploration, beam selection problems are often more effectively modeled using supervised learning approaches when labeled datasets are available. [6]. These RL-based methods demonstrate strong adaptability and robustness in complex systems with uncertainty and faults. However, they typically involve high computational complexity, require iterative

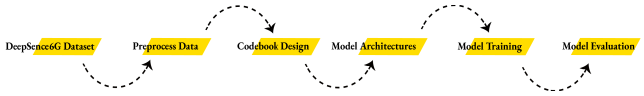

FIGURE 1. Flowchart of proposed methodology.

online exploration, and suffer from slow convergence, which limits their practicality in real-time beam alignment for mmWave communication. In contrast, our work adopts a supervised learning paradigm, where Gaussian Process regression denoises noisy GPS trajectories and Bi-LSTM/Bi-LSTM+Attention architectures directly learn beam predictions from real-world DeepSense 6G measurements, enabling lightweight, low-latency, and deployment-ready solutions.

III. METHODOLOGY

This section details the system model and explicitly defines the challenge of forecasting beam trajectories. Figure 1 depicts a flowchart summarizing the procedures implemented in this research.

The methodology of this study is organized into several subsections, each focusing on a key stage of the proposed framework. First, the *System Model* introduces the mmWave link setup, defining the base station (BS), user equipment (UE), and the beamforming process. The *Problem Formulation* subsection then specifies the optimization objective, namely predicting the optimal beam index using GPS positions instead of full channel state information. Next, the *Dataset* and *Data Preprocessing* subsections describe the DeepSense6G scenarios, codebook configurations, and the data preparation pipeline, including normalization, augmentation, and batching strategies. The *Framework Architecture* outlines the design of the Bi-LSTM and Bi-LSTM with Attention models, emphasizing their robustness to noisy inputs. To address GPS instability, the *Gaussian Process Smoothing of GPS Trajectories* subsection explains the denoising approach, while the subsequent *Problem Formulation* (G) rigorously models GPS observations as noisy samples from a latent continuous function. The *Hyperparameter Estimation* subsection then details the likelihood-based procedure for tuning GP parameters, followed by the *Inference and Smoothing* subsection, which demonstrates how the trained GP is applied to generate reliable trajectory inputs. Collectively, these subsections establish a coherent pipeline from theoretical modeling and dataset preparation to data smoothing and predictive modeling.

A. SYSTEM MODEL

In this work, we examine a millimeter-wave link in which a base station (BS) outfitted with N antennas serves a single-antenna user equipment (UE). The BS selects one of its M predefined beamforming vectors from the codebook for transmission. We model the propagation channel by the complex vector

$$\mathbf{h}_m \in \mathbb{C}^{N\times 1},$$

which captures the per-antenna amplitude and phase shifts between the BS and the UE. When the BS applies beamformer $\mathbf{f}_m$ to transmit the symbol x, the UE receives

$$y = \mathbf{h}_m^{\mathrm{T}}\mathbf{f}_m x + n, \tag{1}$$

as stated in (1), following the conventional mmWave beamforming framework of [14]. where $\mathbf{h}_m \in \mathbb{C}^{N\times 1}$ denotes the complex channel gain vector, capturing the amplitude attenuation and phase rotation from each BS antenna to the UE. The noise term n is modeled as a zero-mean, circularly symmetric complex Gaussian random variable with variance σ^2, i.e.,

$$n \sim \mathcal{CN}(0, \sigma^2).$$

B. PROBLEM FORMULATION

Using received power as the primary performance metric, the BS selects the beamformer that maximizes the expected squared magnitude of the received signal,

$$P = \mathbb{E}\big[|y|^2\big].$$

Accordingly, the optimal beamformer $f^\star$ is given by

$$f^\star = \arg\max_{f\in\mathcal{F}} \big\|\mathbf{h}^{\mathrm{T}} f\big\|^2. \tag{2}$$

Acquiring exact channel state information $\mathbf{h}$ is often impractical. Instead, we aim to infer the best beam solely from the UE's real-time position. Let

$$\mathbf{g} = \begin{bmatrix} g_{\text{lat}} \\ g_{\text{long}} \end{bmatrix} \in \mathbb{R}^2$$

denote the UE's two-dimensional location (latitude and longitude). Our objective is to predict an estimated beam $\hat{f}$ that maximizes

$$\Pr\big(\hat{f} = f^\star \mid \mathbf{g}\big).$$

C. DATASET

Table 1 summarizes the nine DeepSense6G scenarios evaluated in this paper. In each setup, a quasi-omnidirectional transmitter—mounted on a vehicle and operating at 60 GHz—broadcasts mmWave signals and logs its GPS coordinates. The roadside receiver, implemented as an access point (AP) on the sidewalk, is outfitted with 16 antennas and also works at 60 GHz. It captures the incoming signal using an oversampled codebook of 64 predefined beams and, for every measurement instance, selects the beam that maximizes the received power.

During data collection, the vehicle repeats its runs along the lanes in both directions, continuously sending communication and positioning data to the AP. Figure 2 plots the received power across all 64 beams for sample points in Scenario 6; the highest peak in each plot indicates the optimal beam index. The nominal separation between the AP and the transmitter is roughly 20 m. Detailed descriptions of Testbed 1—including the street-level AP module (unit 1) and

TABLE 1. Dataset summary with 60% Training, 20% Validation, and 20% Testing Split.

Scenario	Time	Samples	Split Samples		
			Train (60%)	Val (20%)	Test (20%)
1	Daytime	2411	1447	482	482
2	Nighttime	2974	1784	595	595
3	Daytime	1487	892	297	297
4	Nighttime	1867	1120	373	373
5	Nighttime	2300	1380	460	460
6	Daytime	915	549	183	183
7	Daytime	854	512	171	171
8	Daytime	4043	2426	809	809
9	Daytime	5964	3578	1193	1193

the vehicle-mounted transmitter (unit 2)—can be found in [19,20].

The initial dataset comprises a codebook of $M = 64$ beam indices. To mitigate computational complexity and minimize beam overlap, the codebook is progressively downsampled to reduced sizes of $M \in \{32, 16, 8\}$ beams. To reduce the original $M = 64$ beam codebook to smaller sizes $M \in \{32, 16, 8\}$, we partition the 64 uniformly spaced steering vectors into M contiguous blocks of equal size. Specifically, each block contains

$$\frac{64}{M}$$

adjacent beams, which are grouped and treated as a single representative class. This grouping preserves uniform angular coverage while reducing the beam-training search space from 64 to M, thereby lowering overhead without introducing directional bias.

D. DATA PREPROCESSING

The efficacy of deep learning algorithms is contingent upon the quality and structure of input data. This section delineates the systematic preprocessing methodology necessary for preparing datasets prior to model implementation.

1) **Feature Normalization.** Each feature channel (latitude, longitude, and each beam's received power) was scaled via min–max normalization to the interval [0, 1] according to the minimum and maximum values observed in the training set. This ensures that all inputs contribute comparably during gradient-based optimization and prevents features with larger numeric ranges from dominating the learning process.
2) **Gaussian Noise Augmentation.** To enhance generalization and mitigate sensitivity to measurement inaccuracies, we applied zero-mean Gaussian noise with standard deviation $\sigma = 0.02$ independently to each input feature within the normalized data sequences. These perturbed samples were dynamically generated during training, allowing the Bi-LSTM to learn from minor spatial fluctuations in both positional and signal strength values.
3) **Sequence Construction.** To facilitate sequential modeling, we organized the data into fixed-size sequences comprising T consecutive measurements. At each timestamp t, a feature vector was constructed by concatenating the GPS coordinates with the received power values across N beams:

$$\mathbf{x}_t = [\text{lat}_t, \text{lon}_t, p_{t,1}, p_{t,2}, \ldots, p_{t,N}]$$

4) **Padding and Masking.** Sequences shorter than T were right-padded with zeros and accompanied by a corresponding length mask. During training, sequences were packed using PyTorch's `pack_padded_sequence` utility so that the Bi-LSTM ignored padded timesteps, while maintaining batch-efficient tensor shapes of (batch_size, T, D), where $D = 2 + N$.
5) **Batch Assembly.** A custom `collate_fn` grouped variable-length sequence samples into mini-batches. Each batch provided:
   - A padded tensor of shape (B, T, D),
   - A vector of true sequence lengths for masking, and
   - The corresponding ground-truth labels.

   This organization enables efficient Bi-LSTM training with support for dynamic sequence lengths and on-the-fly data augmentation.

E. FRAMEWORK ARCHITECTURE

This section presents a novel GPS-based beam prediction framework that addresses three critical limitations in existing position-aided mmWave systems: GPS measurement noise, reliance on synthetic datasets, and suboptimal beam selection strategies. We introduce the first dual-head neural architecture that jointly performs beam candidate classification and link-quality regression using real-world GPS trajectories, distinguishing our approach from prior works that treat beam selection as a simple classification problem by predicting both which beams are viable candidates and their expected performance quality. Our framework incorporates four key technical novelties: (1) Gaussian Process GPS denoising that explicitly models and corrects GPS jitter using principled Bayesian smoothing—the first work to systematically address GPS noise in mmWave beam prediction, (2) a bidirectional LSTM with attention mechanism that adaptively weights historical GPS readings to identify the most informative trajectory segments, contrasting with prior works using simple feedforward networks, (3) a multi-task learning framework where the dual-head design simultaneously optimizes beam candidate selection and power prediction for more informed beam choices, and (4) real-world validation using authentic GPS measurements from the DeepSense 6G dataset rather than synthetic simulations. Unlike existing position-aided beam prediction methods that either use synthetic GPS data not reflecting real-world noise characteristics, apply simple neural networks ignoring temporal dependencies, or perform only beam classification without quality assessment, our integrated approach addresses all these limitations simultaneously. The following subsections detail the mathematical formulation, system model, and algorithmic implementation of this framework.

F. GAUSSIAN PROCESS SMOOTHING OF GPS TRAJECTORIES

To mitigate high-frequency jitter in our logged GPS fixes, we model the observed location stream as noisy samples of an underlying smooth trajectory and apply Gaussian-Process (GP) regression to recover a denoised path. Below we detail the formulation and hyperparameter estimation used on our GPS logs.

G. PROBLEM FORMULATION

Let the sequence of GPS observations be denoted by

$$\{(t_i, \mathbf{y}_i)\}_{i=1}^{N}, \quad \mathbf{y}_i = \begin{bmatrix} y_{i,x} \\ y_{i,y} \end{bmatrix}. \tag{3}$$

Each $\mathbf{y}_i$ denotes the recorded latitude and longitude at time t_i. We model these as noisy samples from a latent continuous trajectory drawn from a Gaussian Process (GP):

$$\begin{aligned} \mathbf{f}(t) &\sim \mathcal{GP}\big(\mathbf{m}(t),\ k(t,t') \otimes \mathbf{I}_2\big), \\ \mathbf{y}_i &= \mathbf{f}(t_i) + \eta_i,\ \ \eta_i \sim \mathcal{N}(\mathbf{0},\ \sigma_\eta^2 \mathbf{I}_2). \end{aligned} \tag{4}$$

1) MEAN FUNCTION

We assume a constant mean function:

$$\mathbf{m}(t) = \mu, \quad \mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \tag{5}$$

where μ_x, μ_y are learnable spatial biases.

2) COVARIANCE KERNEL

To balance smooth trajectory modeling with sensitivity to fine-grained motion, we employ a Matérn 3/2 kernel enhanced with additive white noise:

$$k(t,t') = \sigma_f^2 \left(1 + \frac{\sqrt{3}|t-t'|}{\ell}\right) \exp\left(-\frac{\sqrt{3}|t-t'|}{\ell}\right) + \sigma_n^2 \delta_{tt'}, \tag{6}$$

where ℓ denotes the temporal length scale, σ_f^2 is the signal variance, and σ_n^2 quantifies the measurement noise. The full set of trainable hyperparameters is:

$$\Theta = \left\{\mu_x, \mu_y, \ell, \sigma_f^2, \sigma_n^2\right\}. \tag{7}$$

H. HYPERPARAMETER ESTIMATION

We center the observations as

$$\mathbf{y}_c = \big[\mathbf{y}_1 - \mu,\ \ldots,\ \mathbf{y}_N - \mu\big]^\mathsf{T},$$

and maximize the log-marginal likelihood across both columns:
where $[K_\Theta]_{ij} = k(t_i, t_j)$ per (6). We optimize (8), as shown at the bottom of the next page, via L–BFGS–B with multiple restarts.

I. INFERENCE AND SMOOTHING

Given learned Θ, the posterior mean at query time t_* is

$$\hat{\mathbf{f}}(t_*) = \mu + K_\Theta(t_*, \mathbf{t})\, K_\Theta^{-1}\, \mathbf{y}_c,$$

where

$$K_\Theta(t_*, \mathbf{t}) = \big[k(t_*, t_1),\ \ldots,\ k(t_*, t_N)\big].$$

We then use $\hat{\mathbf{f}}(t_i)$ as our denoised fix at each t_i.

By treating the GPS stream as noisy observations of a latent smooth function and fitting a Matérn–3/2 GP with explicit constant mean, we obtain a Bayesian smoother that (i) exactly preserves stationary behavior and (ii) guards against unwanted high-frequency noise—preparing the data for downstream beam-tracking or control.

IV. PROPOSED DUAL-HEAD LSTM-BASED BEAM PREDICTION

In this section, we present our LSTM-based sequence model with a dual-head output designed to (i) shortlist a small set of candidate beams and (ii) estimate the expected link quality of those candidates. By jointly optimizing classification and regression objectives, the model leverages temporal correlations in GPS trajectories to yield both fast and accurate beam selection. This section is structured into several subsections that detail its design and operation. First, the *Sequence Modeling with Bidirectional LSTM* subsection explains how GPS trajectories are processed as sequential inputs using Bi-LSTM layers to capture both past and future dependencies. The *Dual-Head Output Design* then introduces the bifurcated architecture: a mask head for candidate beam classification and a quality head for link-quality regression. Next, the *Joint Loss and Training* subsection defines the multi-task optimization strategy that balances the two objectives, while the *Inference Procedure* illustrates the step-by-step process of shortlisting candidate beams and selecting the optimal one.

A. SEQUENCE MODELING WITH BIDIRECTIONAL LSTM

Let $\{\mathbf{g}_{t-T+1}, \ldots, \mathbf{g}_t\}$ denote the last T GPS readings, where

$$\mathbf{g}_i = \begin{bmatrix} \text{lat}_i \\ \textit{beginalign}*4pt]\text{lon}_i \end{bmatrix}.$$

Each coordinate is min–max normalized using the training-set extrema. These T 2-D vectors form an input tensor of shape $(T, 2)$, which is fed into a single bidirectional LSTM layer with H hidden units per direction. Formally, for each time step i:

$$\overrightarrow{\mathbf{h}}_i = \text{LSTM}_{\rightarrow}\big(\mathbf{g}_i,\ \overrightarrow{\mathbf{h}}_{i-1}\big), \tag{9}$$

$$\overleftarrow{\mathbf{h}}_i = \text{LSTM}_{\leftarrow}\big(\mathbf{g}_i,\ \overleftarrow{\mathbf{h}}_{i+1}\big), \tag{10}$$

and the concatenated hidden state is

$$\mathbf{h}_i = \begin{bmatrix} \overrightarrow{\mathbf{h}}_i \\ \textit{beginalign}*4pt]\overleftarrow{\mathbf{h}}_i \end{bmatrix} \in \mathbb{R}^{2H}.$$

We then apply global average pooling over the T time steps to produce a fixed-length feature vector

$$\mathbf{h} = \frac{1}{T}\sum_{i=1}^{T}\mathbf{h}_i,$$

which captures both past and future trajectory information in a compact embedding.

B. DUAL-HEAD OUTPUT DESIGN

From the shared feature $\mathbf{h}$, the network bifurcates into two task-specific heads:

a: MASK HEAD (MULTI-LABEL CLASSIFICATION)

A lightweight dense block (optional ReLU + dropout) projects $\mathbf{h}$ to M logits, one per beam, followed by a sigmoid activation:

$$\hat{m}_j = \sigma\left(\mathbf{w}_j^\top \mathbf{h} + b_j\right), \quad j = 1, \ldots, M,$$

where $\hat{m}_j \in (0, 1)$ is the probability that beam j belongs to the candidate set. During training, the binary ground-truth mask m_j is obtained by thresholding the measured power,

$$m_j = \begin{cases} 1, & P_j \geq \alpha\, P_{\max}, \\ 0, & \text{otherwise}, \end{cases}$$

and the mask head is optimized with a per-beam binary cross-entropy loss $\mathcal{L}_{\text{mask}}$.

b: QUALITY HEAD (REGRESSION)

In parallel, $\mathbf{h}$ passes through a second dense block to predict a vector of predicted receive powers $\{\hat{q}_j\}_{j=1}^{M}$. We employ a linear output and train with mean squared error against the true measurements $q_j = P_j$:

$$\mathcal{L}_{\text{qual}} = \frac{1}{M}\sum_{j=1}^{M}\left(q_j - \hat{q}_j\right)^2.$$

C. JOINT LOSS AND TRAINING

We combine the two objectives into a single multi-task loss:

$$\mathcal{L} = \lambda_{mask}\,\mathcal{L}_{\text{mask}} + \lambda_{qual}\,\mathcal{L}_{\text{qual}}, \tag{11}$$

where we initially set $\lambda_{mask} = \lambda_{qual} = 1$ and adjust via validation to balance the classification and regression signals. Training is performed with the Adam optimizer (initial learning rate 10^{-2}, decay at epochs 20 and 40), batch size 32, and early stopping on the combined validation loss over 60 epochs.

D. INFERENCE PROCEDURE

At run time, given the most recent T GPS points, the model computes $\{\hat{m}_j\}$ and $\{\hat{q}_j\}$. We first select the top-K beams by descending $\hat{m}_j$, then rank those candidates by predicted quality $\hat{q}_j$, and finally activate the beam

$$j^\star = \arg\max_{j \in \text{candidates}} \hat{q}_j.$$

This two-stage selection minimizes beam sweep overhead while maximizing expected link performance.

V. PROPOSED BI-LSTM WITH ATTENTION AND DUAL-HEAD BEAM PREDICTION

In this section, we introduce our attention-augmented LSTM model with dual-head outputs for efficient and accurate position-aided beam selection. By replacing global average pooling with a learnable temporal weighting mechanism, the network adaptively highlights the most informative GPS readings before jointly performing candidate beam classification and link-quality regression.

The *Attention-Augmented Sequence Embedding* subsection explains how the model adaptively emphasizes the most informative GPS readings, replacing global average pooling with a learnable temporal weighting scheme. Building on this, the *Dual-Head Prediction Architecture* details the two parallel heads: a mask head that classifies candidate beams and a quality head that estimates their link quality. The *Joint Loss and Training Strategy* subsection defines the multi-task optimization objective, while the *Inference and Beam Selection* subsection describes the runtime procedure for selecting the optimal beam. Finally, the *Training Protocol* outlines the experimental setup, including optimizer choices, learning-rate scheduling, and regularization techniques.

A. ATTENTION-AUGMENTED SEQUENCE EMBEDDING

Let $\{\mathbf{g}_{t-T+1}, \ldots, \mathbf{g}_t\}$ be the most recent T GPS coordinates, where

$$\mathbf{g}_i = \begin{bmatrix} \text{lat}_i \\ \textit{beginalign}{*}4pt]\text{lon}_i \end{bmatrix}.$$

After min–max normalization, this $(T \times 2)$ sequence is processed by a single bidirectional LSTM layer with H units per direction, yielding hidden states

$$\overrightarrow{\mathbf{h}}_i = \text{LSTM}_{\rightarrow}(\mathbf{g}_i, \overrightarrow{\mathbf{h}}_{i-1}), \tag{12}$$

$$\overleftarrow{\mathbf{h}}_i = \text{LSTM}_{\leftarrow}(\mathbf{g}_i, \overleftarrow{\mathbf{h}}_{i+1}), \tag{13}$$

$$\mathbf{h}_i = \begin{bmatrix} \overrightarrow{\mathbf{h}}_i \\ \textit{beginalign}{*}4pt]\overleftarrow{\mathbf{h}}_i \end{bmatrix} \in \mathbb{R}^{2H}, \quad i = 1, \ldots, T. \tag{14}$$

$$\log p(\mathbf{y}_c \mid \Theta) = -\tfrac{1}{2}\,\mathbf{y}_c^{\mathsf{T}}\left(K_\Theta\right)^{-1}\mathbf{y}_c - \tfrac{1}{2}\log\left|K_\Theta\right| - \tfrac{N}{2}\log 2\pi, \tag{8}$$

To focus on the most salient time steps, we introduce an attention mechanism of dimension A = H:

$$e_i = v^\top \tanh\big(W\, h_i + b\big) \in \mathbb{R}, \tag{15}$$

$$a_i = \frac{\exp(e_i)}{\sum_{j=1}^{T} \exp(e_j)}, \tag{16}$$

$$c = \sum_{i=1}^{T} a_i\, h_i \in \mathbb{R}^{2H}. \tag{17}$$

The context vector c replaces global average pooling, enabling dynamic emphasis on trajectory segments critical for beam decisions.

B. DUAL-HEAD PREDICTION ARCHITECTURE

From the shared context **c**, the network bifurcates into two parallel heads:

a: MASK HEAD (CANDIDATE BEAM CLASSIFICATION)

- Dense(256) → ReLU → Dropout(0.1)
- Dense(128) → ReLU → Dropout(0.1)
- Output layer: M units → Sigmoid

Each sigmoid output $\hat{m}_j \in (0, 1)$ indicates the probability that beam j belongs in the candidate set.

b: QUALITY HEAD (LINK-QUALITY REGRESSION)

- Dense(256) → ReLU → Dropout(0.1)
- Dense(128) → ReLU → Dropout(0.1)
- Output layer: M units → Linear

Each linear output $\hat{q}_j$ predicts the expected receive-power (or SNR) for beam j.

C. JOINT LOSS AND TRAINING STRATEGY

We optimize both tasks simultaneously via a weighted multi-task loss:

$$\mathcal{L} = \lambda_{\text{mask}}\, \mathcal{L}_{\text{mask}} + \lambda_{\text{qual}}\, \mathcal{L}_{\text{qual}}, \tag{18}$$

where

$$\mathcal{L}_{\text{mask}} = -\frac{1}{M} \sum_{j=1}^{M} \big[m_j \log \hat{m}_j + (1 - m_j) \log(1 - \hat{m}_j)\big],$$

$$\mathcal{L}_{\text{qual}} = \frac{1}{M} \sum_{j=1}^{M} (q_j - \hat{q}_j)^2,$$

and $\lambda_{\text{mask}} = \lambda_{\text{qual}} = 1$ initially. We train using Adam (initial LR $= 10^{-2}$, decay by 0.2 at epochs 20 & 40), batch size 32, with early stopping on the combined validation loss over 60 epochs.

D. INFERENCE AND BEAM SELECTION

At run time:

1) **Feature Extraction:** Last T GPS points → Bi-LSTM → attention → context **c**.
2) **Candidate Shortlisting:** Compute $\{\hat{m}_j\}$, sort descending, select top-K beams.
3) **Quality Ranking:** Evaluate $\{\hat{q}_j\}$ for candidates, choose
$$j^\star = \arg \max_{j \in \text{candidates}} \hat{q}_j.$$
4) **Beam Activation:** Point antenna along beam $j^\star$, avoiding a full M-beam sweep.

This dual-head attention-augmented model jointly minimizes beam-training overhead and maximizes link-quality prediction accuracy.

E. TRAINING PROTOCOL

We train both proposed models end-to-end by minimizing a multi-task loss with the Adam optimizer for up to 60 epochs. The dataset is split into three disjoint subsets: train (60 %), validation (20 %), and test (20 %).

Each sample at time t comprises

- a GPS sequence $\{\mathbf{g}_{t-T+1}, \ldots, \mathbf{g}_t\}$, where $\mathbf{g}_i = [\text{lat}_i, \text{lon}_i]^\top$,
- a power vector $\mathbf{p}_t \in \mathbb{R}^M$.

All coordinates and power values are min–max normalized using the extrema of the training set. During training, we add Gaussian noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, 0.02^2\mathbf{I})$ to the normalized GPS inputs:

$$\widetilde{\mathbf{g}}_i = \widehat{\mathbf{g}}_i + \mathbf{n}.$$

The total loss is

$$\mathcal{L} = \mathcal{L}_{\text{mask}} + \mathcal{L}_{\text{qual}}, \tag{19}$$

with

$$\mathcal{L}_{\text{mask}} = -\frac{1}{M} \sum_{j=1}^{M} \Big[m_j \log \hat{m}_j + (1 - m_j) \log(1 - \hat{m}_j)\Big], \tag{20}$$

$$\mathcal{L}_{\text{qual}} = \frac{1}{M} \sum_{j=1}^{M} (q_j - \hat{q}_j)^2. \tag{21}$$

Here

$$m_j = \mathbb{I}\{p_{t,j} \geq \alpha \max_k p_{t,k}\}, \quad q_j = p_{t,j}.$$

a: OPTIMIZATION

- **Optimizer:** Adam with initial learning rate $\eta_0 = 1 \times 10^{-3}$.
- **LR Schedule:** Multiply η by 0.2 at epochs 20 and 40.
- **Batch Size:** 32.
- **Gradient Clipping:** Clip global norm to 5.
- **Weight Decay:** 1×10^{-5}.
- **Dropout:** Rate 0.1 after each dense layer.
- **Early Stopping:** Stop if validation loss does not improve for 10 consecutive epochs.

F. PARAMETER SELECTION METHODOLOGY

Table 2 summarizes all architectural and training parameters used in our experiments. The following methodology describes how these values were selected:

TABLE 2. Model architecture and training parameters.

Parameter	Value	Selection Method
Bi-LSTM Architecture		
Hidden units (H)	256	Grid search
LSTM layers	1 (bidir.)	Sufficient for GPS
Sequence length (T)	20	2s at 10 Hz
Input dimension	2	GPS (lat, lon)
Attention dim. (A)	256	Equal to H
Dual-Head Layers		
Mask dense 1	256	Grid search
Mask dense 2	128	Dim. reduction
Quality dense 1	256	Match mask head
Quality dense 2	128	Dim. reduction
Output dim.	M	Codebook size
Training		
Optimizer	Adam	Standard DL
Learning rate	10^{-3}	Grid search
LR decay	0.2	At epochs 20, 40
Batch size	32	Memory trade-off
Max epochs	60	Convergence
Early stop	10	Prevent overfit
Gradient clip	5.0	Stability
Weight decay	10^{-5}	L2 regularization
Dropout	0.1	After dense layers
Loss Configuration		
λ_{mask}	1.0	Classification weight
λ_{qual}	1.0	Regression weight
Mask threshold (α)	0.7	70% max power
Data		
GPS noise (σ)	0.02	Measurement error
Train/Val/Test	60/20/20%	Standard split
Gaussian Process		
Mean function	Constant.	Learned (MLE)
Kernel	Matérn 3/2	Smooth-flexible
Length scale	8.7 m	L-BFGS-B
Signal var.	0.42	L-BFGS-B
Noise var.	0.003	L-BFGS-B
Restarts	10	Random init.

1) **Architecture parameters**: Hidden dimensions were selected through grid search over $H \in \{128, 256, 512\}$ on the validation set, with $H = 256$ providing the best trade-off between model capacity and computational efficiency for the 64-beam prediction task.
2) **Sequence length**: Set to $T = 20$ time steps based on the GPS sampling rate (10 Hz) and vehicular coherence time analysis, representing 2 seconds of trajectory history.
3) **Training hyperparameters**: Learning rate was selected via grid search from $\{10^{-2}, 10^{-3}, 10^{-4}\}$, with 10^{-3} showing optimal convergence. Batch size of 32 balanced GPU memory constraints with training stability.
4) **Mask threshold**: The value $\alpha = 0.7$ was empirically determined using the validation set to balance candidate beam set size with prediction accuracy, ensuring the main lobe and strong side lobes are captured while filtering weak beams.
5) **Gaussian Process hyperparameters**: Optimized per scenario using L-BFGS-B with 10 restarts; typical values: $\ell \in [5, 15]$ m, $\sigma_f^2 \in [0.1, 1.0]$, $\sigma_n^2 \in [0.001, 0.01]$.

VI. RESULTS

This section details the experimental evaluation conducted using the proposed models on real-world data encompassing DeepSense Scenarios 1 through 9. GP-BLA (Gaussian Process - Bidirectional LSTM with Attention) leverages the GP's principled smoothing to deliver reliable position estimates, while the Bi-LSTM+Attention module captures temporal dependencies and focuses on the most informative time-steps to predict both the optimal beam index and the expected link quality in a single, lightweight network. Model's performance is assessed using several metrics, including Top-k accuracy, the confusion matrix, and receiver operating characteristic (ROC) curves. To underscore the effectiveness of our method, comparative analyses are also performed against established baseline models.

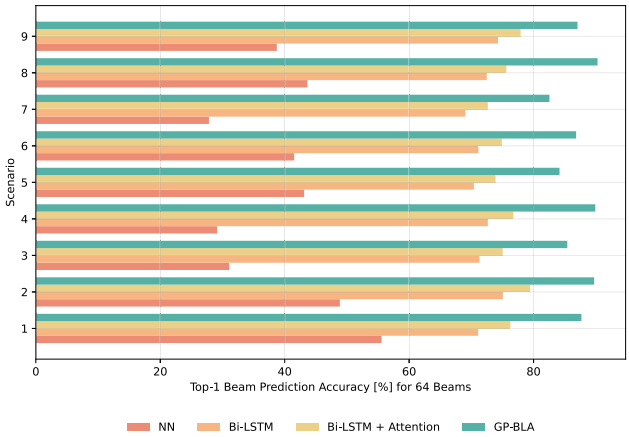

FIGURE 2. Top-1 beam prediction accuracy comparison for 64-beam codebook (M = 64) between the proposed models and the baseline neural network (NN) from [14].

TABLE 3. Top-1 accuracy comparison for 64 beams across scenarios (NN vs. Bi-LSTM vs. Bi-LSTM + Attention vs. GP-BLA).

Scenario	NN	Bi-LSTM	Bi-LSTM-Attn	GP-BLA
1	55.57	71.09	76.27	87.69
2	48.86	75.06	79.45	89.73
3	31.09	71.30	75.04	85.42
4	29.14	72.65	76.75	89.91
5	43.12	70.41	73.88	84.17
6	41.51	71.13	74.90	86.84
7	27.82	69.04	72.65	82.55
8	43.65	72.47	75.63	90.28
9	38.73	74.29	77.92	87.06
Average	**39.9**	**71.9**	**75.72**	**86.41**

A. BASLINE COMPARISON

Figure 2 and Table 3 illustrate the comparison of Top-1 accuracy across all 64 beams for the nine evaluated scenarios. The proposed models demonstrate superior performance over the conventional neural network (NN), with notable gains observed particularly in Scenarios 2, 4, 8, and 9. These results highlight the model's robustness and practical effectiveness in complex, beam-rich real-world environments.

Figure 3 and Table 4 present the Top-1 accuracy results for a reduced set of 32 beams, where the models exhibits even more pronounced performance gains. In particular, Scenarios 1, 2, and 8 show accuracy improvements. These

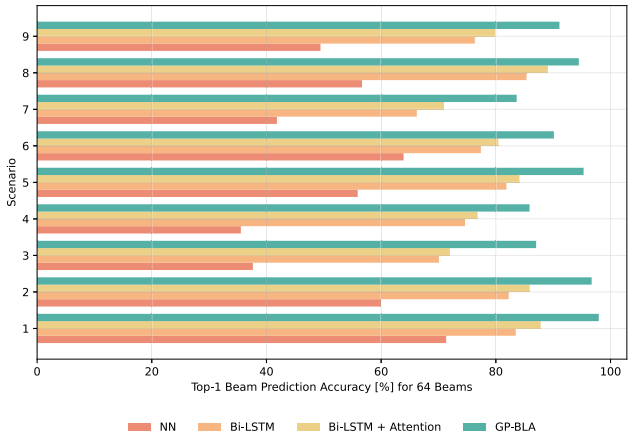

FIGURE 3. Top-1 beam prediction accuracy comparison for 32-beam codebook (M = 32) between the proposed models and the baseline neural network (NN) from [14].

TABLE 4. Top-1 accuracy comparison for 32 beams across scenarios (NN vs. Bi-LSTM vs. Bi-LSTM + Attention vs. GP-BLA).

Scenario	NN	Bi-LSTM	Bi-LSTM-Attn	GP-BLA
1	71.34	83.47	87.83	97.95
2	60.02	82.24	85.91	96.71
3	37.63	70.08	72.00	87.03
4	35.53	74.63	76.82	85.88
5	55.91	81.84	84.12	95.30
6	63.91	77.39	80.44	90.12
7	41.82	66.22	70.95	83.63
8	56.67	85.35	89.08	94.47
9	49.42	76.33	79.92	91.09
Average	**52.7**	**77.1**	**80.7**	**91.8**

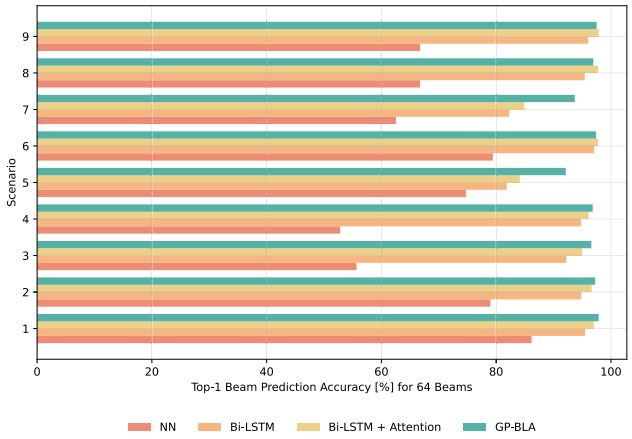

FIGURE 4. Top-1 beam prediction accuracy comparison for 16-beam codebook (M = 16) between the proposed models and the baseline neural network (NN) from [14].

findings affirm the model's strong generalization capability, even under constraints involving a limited number of beams.

When evaluated with a configuration of 16 beams, the models maintain their performance advantage. As shown in Figure 4, and Table 5 the models achieve accuracy levels exceeding 90% across the majority of scenarios, with particularly strong improvements observed in Scenarios 6 and

TABLE 5. Top-1 accuracy comparison for 16 beams across scenarios (NN vs. Bi-LSTM vs. Bi-LSTM + Attention vs. GP-BLA).

Scenario	NN	Bi-LSTM	Bi-LSTM-Attn	GP-BLA
1	86.17	95.51	97.02	97.86
2	78.99	94.82	96.63	97.25
3	55.66	92.24	95.00	96.58
4	52.81	94.77	96.12	96.82
5	74.73	81.84	84.12	92.14
6	79.43	97.04	97.80	97.39
7	62.53	82.29	84.92	93.70
8	66.72	95.43	97.76	96.93
9	66.75	96.03	97.88	97.51
Average	**69.4**	**92.4**	**94.7**	**96.5**

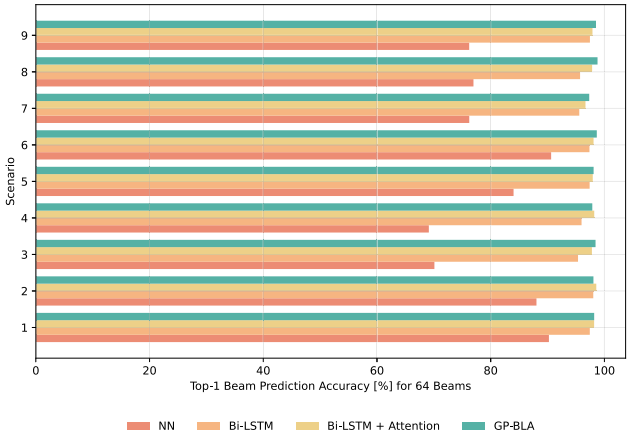

FIGURE 5. Top-1 beam prediction accuracy comparison for 8-beam codebook (M = 8) between the proposed models and the baseline neural network (NN) from [14].

TABLE 6. Top-1 accuracy comparison for 8 beams across scenarios (NN vs. Bi-LSTM vs. Bi-LSTM + Attention vs. GP-BLA).

Scenario	NN	Bi-LSTM	Bi-LSTM-Attn	GP-BLA
1	90.24	97.43	98.19	98.20
2	88.05	98.04	98.57	98.07
3	70.10	95.36	97.80	98.44
4	69.12	95.98	98.21	97.87
5	84.02	97.40	97.95	98.10
6	90.63	97.37	98.09	98.66
7	76.23	95.58	96.67	97.33
8	76.97	95.72	97.84	98.79
9	76.22	97.45	97.90	98.52
Average	**80.1**	**96.6**	**97.9**	**98.3**

9. These results underscore the models effectiveness under moderate beam density conditions.

Figure 5 and Table 6 compare the Top-1 accuracy for 8-beam configuration. The proposed models consistently achieve accuracy exceeding 95% across all scenarios, outperforming the baseline NN [14].

In Scenario 6, the proposed models exhibit robust predictive capability, reliably selecting the optimal beam with high confidence. As illustrated in Figure 6 and Table 7, the models achieves notably high Top-k accuracy, further affirming its effectiveness under this specific scenario. Receiver Operating Characteristic (ROC) curves provide a visual representation of a model's classification performance by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) across varying thresholds for each beam index. Figure 7

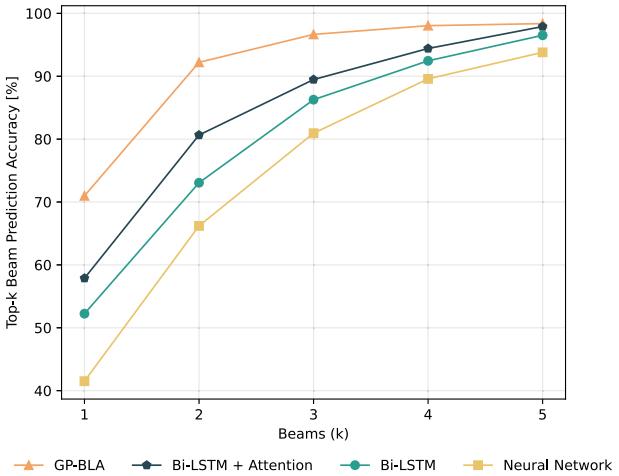

FIGURE 6. Top-k beam prediction accuracy (Top-1 to Top-5) comparison for Scenario 6.

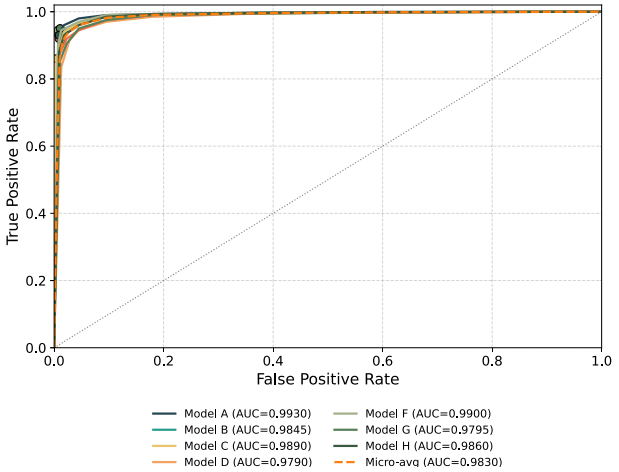

FIGURE 7. Top-k beam prediction accuracy (Top-1 to Top-5) comparison for Scenario 6 using a 8-beam codebook.

illustrates the AUC-ROC curves for $M = 8$ for the first scenario corresponding to the GP-BLA model. A Receiver Operating Characteristic (ROC) curve that lies closer to the top-left corner of the plot—indicating an Area Under the Curve (AUC) value approaching 1—signifies enhanced predictive accuracy of the model. The ROC curves indicate that, for the majority of beam indices, the model achieves a high True Positive Rate (TPR) while maintaining a relatively low False Positive Rate (FPR). This demonstrates the model's effectiveness in distinguishing between correct and incorrect beam selections.

Figure 8 displays the confusion matrix corresponding to the first scenario. The confusion matrix offers a comprehensive assessment of the GP-BLA's model classification performance by enumerating the correct and incorrect predictions associated with each beam index.

1) POWER LOSS

The effectiveness of beam prediction models is most accurately evaluated through the metric of *power loss*, which

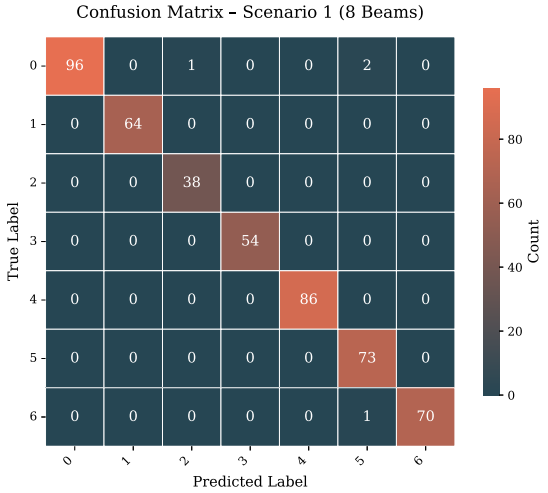

FIGURE 8. Top-k beam prediction accuracy (Top-1 to Top-5) comparison for Scenario 6.

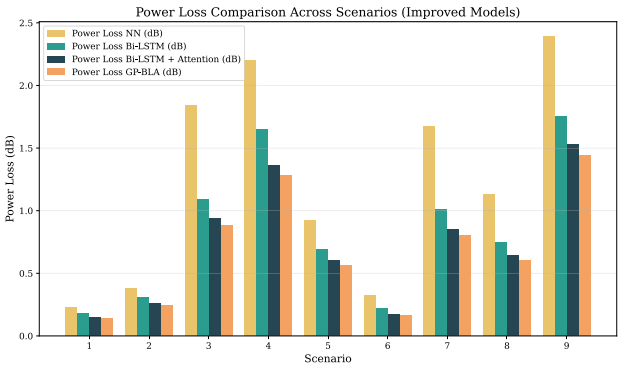

FIGURE 9. Centered average received power, showing power loss and improvements across different scenarios.

TABLE 7. Ranking accuracy evaluation across Top-K predictions for scenario 6 using different models.

Top-K	Neural Network	Bi-LSTM	Bi-LSTM-Attn	GP-BLA
Top-1	41.51	52.24	57.88	70.98
Top-2	66.20	73.06	80.65	92.21
Top-3	80.94	86.27	89.47	96.65
Top-4	89.58	92.45	94.41	98.04
Top-5	93.80	96.51	97.90	98.37

quantifies the degradation in received signal power due to discrepancies between the actual and predicted beam selections. A lower power loss value reflects improved communication quality and more efficient system operation. In the forthcoming analysis, particular emphasis is placed on assessing power loss improvements achieved by the proposed beam prediction models in comparison to existing approaches. The power loss (PL) in decibels is computed using the following expression:

$$\mathrm{PL[dB]} = 10\log_{10}\left(\frac{1}{K}\sum_{k=1}^{K}\frac{P_k^{f^*} - P_n}{P_k^{\hat{f}} - P_n}\right) \tag{22}$$

Here, $P_k^{f^*}$ denotes the received power for the optimal beam corresponding to sample k, $P_k^{\hat{f}}$ represents the received power

TABLE 8. Comparative evaluation of power degradation across beam prediction models in various scenarios.

Scenario	NN (dB)	Bi-LSTM (dB)	Bi-LSTM-Attn (dB)	GP-BLA (dB)
1	0.23	0.18	0.15	0.14
2	0.38	0.31	0.26	0.24
3	1.84	1.09	0.94	0.88
4	2.20	1.65	1.36	1.28
5	0.92	0.69	0.60	0.56
6	0.32	0.22	0.17	0.16
7	1.67	1.01	0.85	0.80
8	1.13	0.75	0.64	0.60
9	2.39	1.75	1.53	1.44

for the predicted beam, P_n is the noise level within the scenario, and K is the total number of evaluated samples. This formulation enables a quantitative comparison between the predicted and ideal beams, with smaller PL values indicating more accurate predictions.

As shown in Figure 9 and Table 8, when applying this metric to the results reported in [14] and comparing them with those achieved by the proposed models, it becomes evident that substantial improvements in power efficiency are observed across all evaluated scenarios.

B. OVERHEAD SAVING

The proposed models deliver improved accuracy in beam prediction while substantially reducing the overhead typically incurred during beam training procedures in millimeter-wave (mmWave) communication systems. This section focuses on quantifying the reduction in training overhead enabled by the position-assisted beam prediction strategy. The overhead savings, denoted as α_{OH}, reflect the proportion of beams that must be trained compared to a full exhaustive search. This metric is inherently influenced by the desired reliability level: higher reliability necessitates evaluating a larger subset of beams, which lowers the potential for overhead reduction. Conversely, tolerating lower reliability permits testing fewer beams, thereby increasing savings. The formal expression for overhead savings is given by:

$$\alpha_{\text{OH}} = 1 - \frac{b}{M}, \tag{23}$$

where b represents the number of beams selected for training, and $M = 64$ denotes the total number of beams in the predefined codebook. Figure 10 illustrates the overhead savings attained by the Bi-LSTM model across representative reliability targets. For a reliability level of 90% (i.e., 10% outage probability), the model achieves beam-training overhead reductions ranging from approximately 82.0% to 95.7% across all nine evaluated scenarios. Even under a more stringent reliability requirement of 99%, the Bi-LSTM-based approach maintains an average overhead reduction of around 67.7%, demonstrating its efficiency in minimizing training overhead while preserving high prediction confidence.

VII. CONCLUSION

This study introduces a unified GPS-enhanced framework combining Gaussian Process regression with Bi-LSTM and Bi-LSTM+Attention architectures for position-based beam prediction in mmWave communication systems, demonstrating Top-1 accuracy of 86.41% and power loss reductions exceeding 1 dB across real-world scenarios. The results are driven by GP-based GPS denoising that addresses measurement noise, combined with dual-head neural architectures that jointly perform beam classification and link-quality regression, converting sequential GPS trajectories into optimal beam predictions. The GP-BLA model outperforms baseline neural networks by 116% in accuracy (86.41% vs 39.9%) while reducing power loss by up to 1.44 dB in challenging scenarios. The GPS denoising approach consistently delivers 8-15% accuracy improvements over raw GPS inputs, while the dual-head architecture design enables simultaneous candidate selection and quality prediction. Similar advantages are observed across all beam configurations (64, 32, 16, 8 beams), with the attention mechanism providing 4-6% additional improvements over standard Bi-LSTM. Practical benefits include up to 95% beam training overhead savings compared to exhaustive search methods while maintaining 90% reliability guarantees. The four stated contributions were realized through: (1) unified framework integration of GP regression with recurrent architectures, (2) advanced GPS denoising proving critical for performance gains, (3) lightweight dual-head design enabling real-time deployment, and (4) comprehensive real-world validation eliminating synthetic dataset dependencies. Future Research Directions: Promising avenues include temporal resolution enhancement and multi-modal sensing, GPS-channel correlation breakdown solutions, multi-user integration challenges, and standardization with beam management protocols.

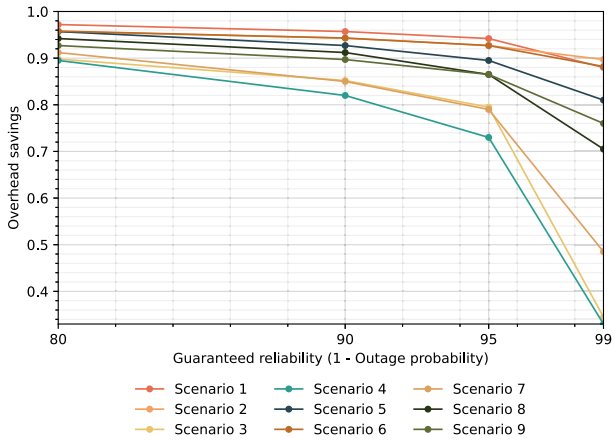

FIGURE 10. Overhead savings plotted against outage probability.

VIII. LIMITATIONS AND FUTURE WORK

The promising results of our GPS-based Bi-LSTM beam prediction framework reveal both its potential and constraints.

A. TEMPORAL RESOLUTION ENHANCEMENT AND MULTI-MODAL SENSING

Our current approach is fundamentally limited by the 10 Hz GPS sampling rate, which creates a temporal mismatch with the mmWave channel coherence time (∼1–10 ms at 60 GHz).

Future work should integrate high-frequency IMUs (100–1000 Hz) to capture sub-second dynamics, particularly during rapid acceleration events where our accuracy drops below 50% at > 3 m/s^2. Additionally, incorporating Doppler radar for velocity estimation and cameras for blockage prediction could address the current blind spot where GPS cannot detect sudden channel obstructions.

B. ADDRESSING GPS-CHANNEL CORRELATION BREAKDOWN

The fundamental assumption that GPS position correlates with optimal beam selection breaks down at 60 GHz, where 1 cm movements cause 30° phase shifts. Future research should explore differential GPS encoding using velocity $[\Delta\text{lat}/\Delta t,\ \Delta\text{lon}/\Delta t]$ and acceleration $[\Delta^2\text{lat}/\Delta t^2,\ \Delta^2\text{lon}/\Delta t^2]$ features rather than absolute positions.

C. MULTI-USER AND NETWORK INTEGRATION CHALLENGES

Extending to multi-user scenarios reveals the hidden computational complexity: $\mathcal{O}(N^2)$ for N users due to inter-beam interference calculations. Future work should develop distributed beam prediction where each UE predicts locally, and the base station performs lightweight conflict resolution. Integration with 5G/6G network slicing requires beam predictions to be QoS-aware, suggesting a multi-task learning framework that jointly optimizes beam selection and resource allocation. The current approach also lacks standardized interfaces with 3GPP beam management procedures (P1, P2, P3), requiring protocol-aware training that respects SSB periodicity and CSI-RS configurations.

ACKNOWLEDGMENT

During the preparation of this work, the author(s) used ChatGPT, an AI language model developed by OpenAI, in order to enhance the clarity, coherence, and readability of the manuscript's writing. After using this tool, they reviewed and edited the content as needed and take full responsibility for the content of the published article.

REFERENCES

[1] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[2] V. A. Nugroho and B. M. Lee, "GPS-aided deep learning for beam prediction and tracking in UAV mmWave communication," *IEEE Access*, vol. 13, pp. 117065–117077, 2025, doi: 10.1109/ACCESS.2025.3586594.

[3] M. Q. Khan, A. Gaber, M. Parvini, P. Schulz, and G. Fettweis, "A low-complexity machine learning design for mmWave beam prediction," *IEEE Wireless Commun. Lett.*, vol. 13, no. 6, pp. 1551–1555, Jun. 2024.

[4] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*, 2014.

[5] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[6] D. D. S. Brilhante, J. C. Manjarres, R. Moreira, L. de Oliveira Veiga, J. F. de Rezende, F. Müller, A. Klautau, L. L. Mendes, and F. A. P. de Figueiredo, "A literature survey on AI-aided beamforming and beam management for 5G and 6G systems," *Sensors*, vol. 23, no. 9, p. 4359, Apr. 2023, doi: 10.3390/s23094359.

[7] R. Wang, C. She, Y. Li, and B. Vucetic, "Attention mechanism for beam prediction in mmWave communications with mission-critical applications," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Rio de Janeiro, Brazil, Dec. 2022, pp. 1454–1459, doi: 10.1109/GCWKSHPS56602.2022.10008556.

[8] K. K. Biliaminu, S. A. Busari, J. Rodriguez, and F. Gil-Castiñeira, "Beam prediction for mmWave V2I communication using ML-based multiclass classification algorithms," *Electronics*, vol. 13, no. 13, p. 2656, Jul. 2024, doi: 10.3390/electronics13132656.

[9] M. Ma, N. T. Nguyen, N. Shlezinger, Y. C. Eldar, and M. Juntti, "Attention-enhanced learning for sensing-assisted long-term beam tracking in mmWave communications," 2025, *arXiv:2509.11725*.

[10] D. Pjanić, G. Tian, A. Reial, X. Cai, B. Bernhardsson, and F. Tufvesson, "Illuminating the path: Attention-assisted beamforming and predictive insights in 5G NR systems," 2025, *arXiv:2505.18160*.

[11] M. Arvinte, M. Tavares, and D. Samardzija, "Beam management in 5G NR using geolocation side information," in *Proc. 53rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2019, pp. 1–6.

[12] S. Rezaie, C. N. Manchón, and E. de Carvalho, "Location- and orientation-aided millimeter wave beam selection using deep learning," in *Proc. ICC - IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1–6, doi: 10.1109/ICC40277.2020.9149272.

[13] Y. Heng and J. G. Andrews, "Machine learning-assisted beam alignment for mmWave systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 4, pp. 1142–1155, Dec. 2021.

[14] 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)*, Mar. 2023, pp. 1824–1829.

[15] C. Slezak, V. Semkin, S. Andreev, Y. Koucheryavy, and S. Rangan, "Empirical effects of dynamic human-body blockage in 60 GHz communications," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 60–66, Dec. 2018.

[16] T. Bai and R. W. Heath Jr., "Coverage analysis for millimeter wave cellular networks with blockage effects," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2013, pp. 727–730.

[17] J. C. Aviles and A. Kouki, "Position-aided mm-wave beam training under NLOS conditions," *IEEE Access*, vol. 4, pp. 8703–8714, 2016.

[18] G. Charan, A. Hredzak, C. Stoddard, B. Berrey, M. Seth, H. Nunez, and A. Alkhateeb, "Towards real-world 6G drone communication: Position and camera aided beam prediction," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Rio de Janeiro, Brazil, Dec. 2022, pp. 2951–2956, doi: 10.1109/GLOBECOM48099.2022.10000718.

[19] M. S. Khan, Q. Sultan, and Y. S. Cho, "Position and machine learning-aided beam prediction and selection technique in millimeter-wave cellular system," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2020, pp. 603–605, doi: 10.1109/ICTC49870.2020.9289508.

[20] M. Fabiani, D. Silva, A. Abdallah, A. Celik, and A. M. Eltawil, "Multi-modal sensing and communication for V2V beam tracking via camera and GPS fusion," in *Proc. 58th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2024, pp. 1–6.

[21] J. Ababneh, H. Attar, A. Solyman, A. Alrosan, and R. Agieb, "Efficient beam selection in mmWave cellular systems using neural networks and K-nearest neighbors based on GPS coordinates," *Appl. Math. Inf. Sci.*, vol. 19, pp. 659–670, Apr. 2025, doi: 10.18576/amis/190314.

[22] J. Gui and Y. Liu, "Enhancing energy efficiency for cellular-assisted vehicular networks by online learning-based mmWave beam selection," *EURASIP J. Wireless Commun. Netw.*, vol. 2022, no. 1, Dec. 2022, Art. no. 1, doi: 10.1186/s13638-021-02080-5.

[23] N. Suneetha and P. Satyanarayana, "Intelligent channel estimation in millimeter wave massive MIMO communication system using hybrid deep learning with heuristic improvement," *Int. J. Commun. Syst.*, vol. 36, no. 5, Mar. 2023, Art. no. e5400, doi: 10.1002/dac.5400.

[24] K. Ma, S. Du, H. Zou, W. Tian, Z. Wang, and S. Chen, "Deep learning assisted mmWave beam prediction for heterogeneous networks: A dual-band fusion approach," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 115–130, Jan. 2023, doi: 10.1109/TCOMM.2022.3222345.

[25] N. Chen, H. Lin, Y. Zhao, L. Huang, X. Du, and M. Guizani, "Low complexity closed-loop strategy for mmWave communication in industrial intelligent systems," *Int. J. Intell. Syst.*, vol. 37, no. 12, pp. 10813–10844, Dec. 2022.

[26] Q. Xue, C. Ji, S. Ma, J. Guo, Y. Xu, Q. Chen, and W. Zhang, "A survey of beam management for mmWave and THz communications towards 6G," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1520–1559, 3rd Quart., 3, 2024, doi: 10.1109/COMST.2024.3361991.
[27] W. Yi, W. Zhiqing, and F. Zhiyong, "Beam training and tracking in mmWave communication: A survey," *China Commun.*, vol. 21, no. 6, pp. 1–22, Jun. 2024, doi: 10.23919/jcc.ea.2021-0873.202401.
[28] Y. Sheng, K. Huang, L. Liang, P. Liu, S. Jin, and G. Y. Li, "Beam prediction based on large language models," *IEEE Wireless Commun. Lett.*, vol. 14, no. 5, pp. 1406–1410, May 2025, doi: 10.1109/LWC.2025.3543567.
[29] C. Zheng, J. He, C. G. Kang, G. Cai, Z. Yu, and M. Debbah, "M2BeamLLM: Multimodal sensing-empowered mmWave beam prediction with large language models," 2025, *arXiv:2506.14532*.
[30] J. Li, L. Yuan, W. Cheng, T. Chai, and F. L. Lewis, "Reinforcement learning for synchronization of heterogeneous multiagent systems by improved Q-functions," *IEEE Trans. Cybern.*, vol. 54, no. 11, pp. 6545–6558, Nov. 2024, doi: 10.1109/TCYB.2024.3440333.
[31] H. Shi, W. Gao, X. Jiang, C. Su, and P. Li, "Two-dimensional model-free Q-learning-based output feedback fault-tolerant control for batch processes," *Comput. Chem. Eng.*, vol. 182, Mar. 2024, Art. no. 108583, doi: 10.1016/j.compchemeng.2024.108583.
[32] D. Tandler, S. Doerner, M. Gauger, and S. ten Brink, "Deep reinforcement learning for mmWave initial beam alignment," in *Proc. 26th Int. ITG Workshop Smart Antennas 13th Conf. Syst., Commun., Coding (WSA SCC)*, Braunschweig, Germany, Feb. 2023, pp. 1–6.
[33] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, "Optimal beam association for high mobility mmWave vehicular networks: Lightweight parallel reinforcement learning approach," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5948–5961, Sep. 2021, doi: 10.1109/TCOMM.2021.3088305.

MUHAMMAD HASEEB HASHIR received the B.S. degree in software engineering from COMSATS University Islamabad, Pakistan. He is currently pursuing the master's degree in information and communication engineering with Yeungnam University, Republic of Korea. He was a Software Developer with the Daxno Technologies (Pvt.) Ltd., Islamabad, Pakistan. His research interests include artificial intelligence, 5G, beyond-5G, and wireless networks.

MEMOONA received the B.S. degree from the Department of Computer Science, COMSATS University Islamabad, Pakistan. She joined the Department of Information and Communication Engineering, College of Engineering, Yeungnam University, Gyeongsan, South Korea, where she is currently pursuing the Ph.D. degree. She was a Web Developer with Cane Technologies, Pakistan. Her research interests include WiFi 7 and artificial intelligence. She is dedicated to contributing to wireless networks.

SUNG WON KIM received the B.S. and M.S. degrees from the Department of Control and Instrumentation Engineering, Seoul National University, South Korea, and the Ph.D. degree from the School of Electrical Engineering and Computer Sciences, Seoul National University. He was a Researcher with the Research and Development Center, LG Electronics, South Korea. He was a Postdoctoral Researcher with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. He joined the School of Computer Science and Engineering, Yeungnam University, Gyeongsangbuk-do, South Korea, where he is currently a Professor. His research interests include resource management, wireless networks, mobile computing, performance evaluation, and machine learning.