International Journal of Communication Systems

WILEY

RESEARCH ARTICLE

Adaptive Deep Learning Technique With Deep Feature Extraction for Accurate Path Loss Estimation in Millimeter-Wave Wireless Communication Environments

B. M. R. Manasa[1] | Vijayakumar Kondepogu[1] | Ch V Ravi Sankar[1] | P. Sankara Rao[2] | A. Lakshmi Narayana[1]

[1]Department of Electronics and Communication Engineering, Aditya University, Surampalem, Andhra Pradesh, India | [2]Department of Electronics and Communication Engineering, Avanthi Institute of Engineering and Technology, Vizianagaram, Andhra Pradesh, India

Correspondence: B. M. R. Manasa (manasabmr@adityauniversity.in)

Received: 4 July 2025 | **Revised:** 5 November 2025 | **Accepted:** 22 December 2025

Keywords: Adaptive Residual Bidirectional Gated Recurrent Unit | millimeter-wave wireless communication | path loss estimation | Pyramid Multihead Convolutional Cross Attention Network | Updated Random Attribute–based Sculptor Optimization

ABSTRACT

Millimeter-wave (mmWave) communication plays a crucial role in wireless systems due to its high data rate capabilities and suitability for 5th generation (5G) networks. However, mmWave signals confront significant propagation issues, which include greater path loss, major attenuation from blockages, and sparse multipath propagation, constraining coverage and consistency. Accurate path loss validation is a serious and composite task for successful network planning, optimization, and resource allocations. To overcome these limitations, effective deep learning–based path loss estimation in mmWave communication systems is developed in this research work. Initially, the required data are collected from the standard datasets and given to the preprocessing phase. Once the data are preprocessed, they are given into the deep feature extraction phase, and it is done by applying the Pyramid Multihead Convolutional Cross Attention Network (PMC-CANet). The ability to ensure the efficiency of next-generation wireless networks is what makes it effective in feature extraction tasks. Finally, the path loss estimation process is performed on the extracted deep features through Adaptive Residual Bidirectional Gated Recurrent Unit (AR-BiGRU), where several parameters are tuned using the Updated Random Attribute–based Sculptor Optimization (URA-SO). One of the primary advantages of using AR-BiGRU with USOA for path loss estimation is its ability to process large, high-dimensional datasets, which can include not only geographical and environmental information but also temporal data, such as time-of-day or seasonal variations in path loss. The optimal solution outcome can be achieved by using the developed model. Then, its effectiveness is validated by comparing it with other existing models. This proposed system provides a consistent and best solution for tackling the problems of mmWave signal attenuation, thus enhancing the effectiveness and performance of next-generation wireless networks. The outcomes reveal that the proposed URA-SO-AR-BiGRU obtained an accuracy of 97.12% when taking the batch size as 15, leading to highly reliable and precise path loss estimations.

1 | Introduction

Millimeter waves (mmWaves) are characterized by frequencies between 30 and 300 GHz that are integral to next-generation wireless technologies, which include high-speed data transmission, satellite systems, and 5G networks [1]. The mmWave signals have broader bandwidth, and they provide higher data rates when compared to lower frequency counterparts. However, path loss is one of the main drawbacks of mmWave communication [2]. The reduction in signal strength caused by the material interactions, weather conditions, environmental obstructions, and distance is known as path loss [3].

Compared to lower frequency signals, mmWave signals are prone to physical barriers, degradation, and humidity issues [4]. In order to maximize spectral utilization, ensure consistent connectivity and maintain network efficiency, a path loss prediction model is essential [5]. Accurately determining path loss is a key component of wireless communication, as it quantifies signal attenuation over propagation distance [6]. As the deployment of mmWave is increasing in 5G infrastructures, the path loss issue also increases [7]. Therefore, advanced tools for adaptive and precise path loss modeling are needed to handle the signal fluctuations for better resource management across varying operational conditions [8].

Traditional approaches to estimate path loss in mmWave communication systems are developed based on mathematical formulations that forecast signal degradation by considering different attributes like environment and propagation behavior [9]. Techniques like free-space path loss (FSPL) are suitable for line-of-sight conditions, as they correlate distance and signal degradation [10]. However, this method is difficult to implement in complex environments [11]. More advanced methods, such as the log-distance model, are used for finding the obstacles that lead to path loss by introducing a path loss exponent, but still, this technique relies on generalized parameters [12].

Conventional path loss estimation techniques in mmWave communications face several limitations that affect their accuracy and adaptability [13]. Traditional methods use static parameters, so they do not estimate the path loss by analyzing factors such as physical obstructions and atmospheric changes [14]. Additionally, the unpredictable nature of multipath propagation including scattering, diffractions, and signal reflections is not effectively handled by the classical models, which results in reduced estimation precision during path loss [15]. Mobility is also one of the issues in path loss prediction as vehicular networks or mobile communications demand real-time adaptability [16]. Moreover, the CI and FI rely heavily on extensive measurement datasets that limit their scalability and transferability across varying environments [17]. To address these issues, the present work introduces a deep learning-based approach for mmWave path loss estimation by overcoming the limitations of traditional techniques.

The unique contributions of the presented framework are listed below:

- To design an effective and intelligent path loss estimation model that helps to increase signal strength by identifying the cause of path loss issues. The data transmission via next-generation wireless communication is enhanced by the developed model. Despite complex environmental dynamics and propagation conditions, this model still ensures precise path loss prediction. The proposed path loss determination mechanism is beneficial in enhancing network reliability and improving signal planning efficiency for offering feasible communication.
- To employ a Pyramid Multihead Convolutional Cross Attention Network (PMC-CANet) model for deep feature extraction from the preprocessed data. This model consists of pyramid-based multihead convolutions for learning the local and global patterns in the pre-processed data. The multihead cross attention is useful in emphasizing the contextually significant features. The PMC-CANet approach accurately handles nonlinear characteristics in the data and achieves high discriminative capability by combining pyramid convolutional and multihead cross attention. Feature variability issues in data captured from the mm-Wave environment are handled by this network.
- To develop an Adaptive Residual Bidirectional Gated Recurrent Unit (AR-BiGRU) model for performing accurate path loss estimation in mmWave communication. The development model's use of bidirectional sequence learning and residual learning assists in capturing the past and future dependencies of temporal data. Moreover, the degradation issues are solved by the residual connections in the proposed network. Stable gradient flow is achieved by the AR-BiGRU network for processing a huge volume of features. Highly reliable and context-aware path loss prediction outcomes are attained by the suggested model due to its ability to integrate the temporal, spatial, and environmental factors.
- To introduce the Updated Random Attribute–based Sculptor Optimization (URA-SO) algorithm for tuning the hyperparameters of the AR-BiGRU model. The faster convergence as well as enhanced search efficiency of the URA-SO algorithm helps to optimize the important parameters for obtaining accurate path loss determination results. The URA-SO algorithm is being introduced to adjust the hyperparameters of the AR-BiGRU model.

The logical flow of this research work is organized as follows: Section 2 gives the existing research works along with the problem formulation of prior models. Section 3 outlines the significance of path loss prediction in mm-Wave communication and provides an overview of the developed model. Section 4 explains the preprocessing strategy and the details of the feature extraction procedure, along with the optimization approach. Section 5 explains the deep learning-based estimation process using the proposed network and elaborates its objective function. The performance evaluation outcomes are discussed in Section 6, and the conclusion of this work is given in Section 7.

2 | Literature Survey

2.1 | Related Works

In 2024, Afape et al. [18] have introduced a machine learning model designed for path loss prediction known as the Automated Hyperparameter-Tuned Stacking Ensemble Regression (AHT-SEML) framework. Automated hyperparameter tuning techniques were used to optimize the performance of a meta-regressor and a set of base regressors in this model. To validate the effectiveness of this approach, the researchers employed path loss data derived from a composite 3D ray-tracing image-method model and applied it to various scenarios.

In 2024, Zakeri et al. [19] have suggested a hybrid path loss estimation approach for indoor environments with the help of machine learning techniques and electromagnetic

calculations. After obtaining real data on path loss, they compared traditional empirical models with advanced models such as cat boosting and deep neural networks. The machine learning–based models yielded better outcomes as they are capable of modeling complex signal behavior in indoor wireless environments.

In 2024, Brata and Zakia [20] have tackled path loss prediction using limited satellite imagery. The authors utilized convolutional neural network (CNN)–based networks to predict path loss. Among the two different models, the visual geometry group (VGG-16) offered higher accuracy. Their models surpassed the 3GPP Urban Macro (UMa) empirical baseline by reducing the error rates.

In 2024, Sung et al. [21] have developed a deep learning-based model for predicting mmWave path loss in vehicle-to-Vehicle (V2V) communications under dynamic and poor road conditions. They trained their model on a simulated environment for handling real-world road and weather variability. Validation of the effectiveness of deep learning for accurate path loss forecasting in vehicular networks was provided by the results.

In 2024, Thrane et al. [22] have analyzed the ability of conventional channel models and a deep learning-based model for the path loss estimation process by utilizing satellite images. They trained their model and evaluated its performance against ray-tracing models as well as conventional models. The frequency of predictions was taken into account when their deep learning approach improved performance and provided effective results in unseen locations.

In 2024, Pawar et and Venkatesan [23] have developed a path loss prediction strategy using CNNs, and the prediction errors were minimized using Particle Swarm Optimization (PSO). An outdoor simulation scenario was considered in this study using DeepMIMO datasets. Their CNN model was able to achieve high prediction accuracy and reduce RMSE values. Therefore, the suggested model was capable of handling multifrequency and blockage-prone outdoor environments.

In 2024, Kayaalp et al. [24] have introduced a deep learning-based path loss prediction framework for various vegetative and coastal terrains using 5G frequency bands. They applied the recurrent neural network (RNN) and long short-term memory (LSTM) methods for the estimation process and identified that RNN provided better predictive accuracy across different terrains. According to their model, coastal environments experience greater path loss than vegetation regions.

In 2024, Cheng et al. [25] have developed a deep learning model for 5G suburban path loss prediction by incorporating attention mechanisms and dilated convolutions. The attention-enhanced CNN (AE-CNN) architecture extracted comprehensive features using global context blocks and distance-embedded inputs. The strength in capturing complex propagation characteristics is evident in the AE-CNN's lower RMSE score compared to conventional empirical models.

In 2025, Hadji and Nedil [26] have proposed a deep neural network (DNN)–based approach to forecast path loss by utilizing massive multiple-input multiple-output (MIMO) propagation measurements at 28 GHz in an underground mine setting. Evaluated with classical empirical frameworks, including log-distance and multislope techniques, the developed DNN system reveals higher predictive accuracy, gaining an RMSE of 1.46 dB and a correlation of 99.17%, successfully extracting the intricate propagation characteristics of the environment.

In 2025, Hussain [27] has developed a UNet-based deep learning model that integrates multiscale feature extraction, convolution-based feature fusion, and an atrous spatial pyramid pooling (ASPP) bottleneck for efficient context aggregation. The model forecasts pathloss maps from log-distance, line-of-sight (LOS) mask, and building mask inputs. The experimental results show that the proposed model performs better compared to other traditional models.

In 2025, Robinson et al. [28] have suggested a machine learning model to introduce an optimized path loss forecasting framework for urban drive 5G network settings. At the training phase, various IoT-based data such as received power, path loss, bandwidth, frequency, and distance were used. Then, ensemble machine learning classifiers were validated and attained scalable performance.

2.2 | Problem Statement

The mmWave communication is essential for next-generation wireless networks, enabling higher data rates and lower latency. However, estimating path loss accurately poses significant challenges that impact network performance and reliability. Environmental factors such as obstacles, atmospheric conditions, and material properties can cause path loss to vary, which affects signal propagation. Additionally, traditional path loss models often fail to account for the complexity of mmWave interactions, leading to inaccurate predictions that can degrade system efficiency. Effective analysis requires advanced tools to manage and process the vast amounts of data generated by modern measurement systems, which is another challenge. For addressing these issues, a deep learning–aided path loss estimation mechanism is suggested that leveraging advanced neural networks to enhance accuracy and optimize wireless communication performance is developed in this work. Features and challenges of the traditional path loss estimation models are given in Table 1, and the following section provides the research gaps:

- The accurate modeling of mobility's impact on signal propagation is a challenge when estimating path loss for mmWave communications. In dynamic environments, such as vehicular networks or mobile device usage, signal characteristics change rapidly due to movement, leading to unpredictable variations in path loss. One way to address this issue is through deep learning models. These models process sequential path loss data, capturing trends influenced by movement, speed, and direction. Deep learning's continuous updating of predictions using real-time data allows for adaptive and precise path loss estimation, which improves connectivity in dynamic wireless environments.

TABLE 1 | Advancements and limitations of deep learning-based path loss estimation models for mmWave.

Author [citation]	Methodology	Features	Challenges
Afape et al. [18]	SEML-AHT	It boosts prediction accuracy by combining multiple base regressors and also ensuring reliable path loss estimation across diverse conditions. It achieves lower error rates compared to traditional models, making it a more effective tool for mmWave system path loss prediction.	It relies on simulated path loss data, which may not fully capture real-world complexities. At present, it is only applicable to mmWave systems in various cities.
Zakeri et al. [19]	Deep Neural Network	It ensures precise path loss estimation by using different polarization settings, enhancing accuracy for NLOS and LOS cases. The reliable performance across various environments and frequencies is provided by its strong design and simplicity.	It depends on hyperparameter tuning, which can be slow and resource-intensive, making real-time use challenging. It enhances accuracy but needs a large dataset to work well in different indoor settings.
Brata and Zakia [20]	VGG-16	It improves accuracy without costly data collection. It has lower RMSE and MAE, making it more reliable for small datasets	It does not consider transmitter-receiver distance, which may affect accuracy. It predicts faster but has higher errors, reducing reliability.
Sung et al. [21]	Deep learning	It considers weather, channels, and traffic, improving path loss prediction in real-world driving conditions. It reduces measurement costs and saves time by using deep learning instead of traditional modeling methods.	It relies on simulation data, which may not fully reflect real-world variations.
Thrane et al. [22]	DL-based ray-tracing	It uses satellite images that help neural networks make more accurate path loss predictions. They provide important details like buildings and trees that affect signal strength.	It requires extra resources for processing satellite images. The accuracy is dependent on the availability of good satellite images, which may not always be available.
Pawar and Venkatesan [23]	PSO	It predicts path loss more precisely with a lower error rate. It considers multiple parameters like users, base stations, and user positions, improving predictions.	It requires significant processing power to train and run effectively. Its accuracy depends on the quality of numerical data used.
Kayaalp et al. [24]	RNN	Its model provides better accuracy in predicting path loss, especially in coastal and vegetation environments. Model performance is improved by optimizing training using K-fold validation and hyperparameter tuning.	Its effectiveness depends on the dataset structure, limiting its general applicability. It does not account for extreme weather conditions like snow and rain.
Cheng et al. [25]	AE-CNN	It efficiently predicts path loss for 28-GHz mmWave in suburban areas. It simplifies preprocessing and input generation, making the application easier.	Its accuracy depends on environmental factors like buildings and streets, and it does not consider moving vehicles. Improving performance necessitates more tuning and optimization.

- The challenge of inaccurate path loss estimation in conventional models can significantly impact wireless communication efficiency. Environmental variability is often overlooked by traditional methods, resulting in imprecise predictions that impact signal reliability and network performance. This limitation can result in suboptimal resource allocation, increased interference, and degraded connectivity. One way to address this issue is through advanced data preprocessing techniques. By cleaning, normalizing, and structuring large datasets of signal measurements, data preprocessing enhances the accuracy of path loss models.

- Diffraction and scattering effects can cause distortion of signal propagation, which is another significant challenge. Due to the short wavelength of mmWave, signals interact strongly with surface textures and small obstacles, leading to unpredictable variations in path loss measurements. This challenge makes it difficult to build accurate models that can generalize across diverse urban and indoor environments. Deep feature extraction techniques can be used to address this issue. By leveraging deep feature extraction models, they automatically identify complex signal behaviors influenced by diffraction and scattering.
- Distinguishing between different propagation environments, such as urban, indoor, and rural settings, is a significant challenge. Each environment has unique characteristics that affect signal attenuation, reflection, and diffraction, making it challenging to apply a single path loss model across all scenarios. One way to address this issue is through residual connection. Residual can categorize propagation conditions automatically by training deep learning models on labeled path loss measurements from different environments.

3 | Detailed Representation of Proposed Accurate Path Loss Estimation Model in mm-Wave Wireless Communication Environments

3.1 | Significance of Predicting Path Loss in mm-Wave Wireless Communication Environments

As mm-Wave transmissions use high-frequency signals for sharing data, they are often affected by path loss issues. Path loss problems may arise due to various environmental factors, such as surface reflections, air absorption, and physical barriers. When compared to communication technologies that use lower frequency signals, mm-wave transmissions are highly susceptible to small-scale environmental changes [26]. In order to guarantee better signal coverage, minimize communication interruptions, and preserve overall system efficiency, precise path loss prediction is needed. Furthermore, mm-wave communication systems are widely operated in crowded urban areas or intricate indoor spaces, but the propagation behavior might differ over short distances in these places. The use of building materials, vegetation, transportation patterns, and time-based variations like seasonal shifts can also lead to path loss issues during mm-wave transmissions. To achieve ultralow latency and high data throughput in mm-Wave communication, accurate path loss prediction is useful, as it has a direct impact on network planning, resource allocation, and beamforming techniques. The use of mmWave wireless communication systems in designing and optimizing next-generation networks like 5G to predict path loss has increased in recent days. Most path loss models consider static factors, but these approaches might not perform well in identifying path loss by analyzing physical barriers. The process of accurately forecasting path loss helps to improve user experience and quality of service (QoS) and enhances the long-term scalability of mm-Wave wireless communication networks.

3.2 | Overall View of Proposed Path Loss Estimation Model

An advanced deep learning–based model for accurate path loss estimation in mm-Wave wireless communication environments is proposed in this research. This model is capable of overcoming frequent issues like nonlinear propagation behaviors, high-dimensional input data, and dynamic environmental conditions. Standardized data are collected to estimate the path loss. Then, preprocessing is conducted on the input data for cleaning and normalizing the input to increase the accuracy of forecasting. This preprocessing procedure helps to lower the noise and bias issues. Deep features are extracted by the suggested PMC-CANet with preprocessed data. The multihead cross-attention procedure helps to capture the relevant features. Different aspects of the preprocessing data that are useful for the prediction process are efficiently retrieved in this stage. Feature retrieval is done by the pyramid convolution module in the PMC-CANet model at different scales. After the extraction of deep features, the AR-BiGRU model is employed to determine the path loss. Accurate path loss estimation in mm-Wave transmissions can be achieved with the BiGRU approach's dual capability of learning sequential dependencies in both forward and backward temporal directions. The proposed model is capable of handling shifting environmental circumstances and obtains better prediction outcomes. The complexity issues and vanishing gradient problems are managed by the residual connection incorporated with the Bi-GRU model. Additionally, the URA-SO algorithm is used to modify different parameters in the AR-BiGRU for enhancing the estimation accuracy. Through the optimization process, the AR-BiGRU mechanism's learning effectiveness is enhanced. The ability of AR-BiGRU to handle diverse datasets with spatial, environmental, and temporal characteristics is strengthened by the combination URA-SO algorithm. The suggested framework's accuracy and scalability are confirmed through experimental analysis using different existing techniques. The architectural framework of the proposed path loss estimation model in mm-Wave communication is exhibited in Figure 1.

3.3 | Dataset Description

In this work, the dataset was derived from simulated path loss data created using a MATLAB R2023a–based Composite 3D Raytracing-Image (CRTI) propagation framework [18] taken online (from https://doi.org/10.1016/j.rineng.2024.102289). These data consist of environmental and structural attributes, such as atmospheric variations, architectural styles, building heights and constructions, topography, and overall urban layouts of different locations. The outdoor mmWave signal characteristics at frequencies of 28, 38, 60, and 73 GHz are gathered in this database from four urban regions. The gathered samples are termed as C_w. Total data are

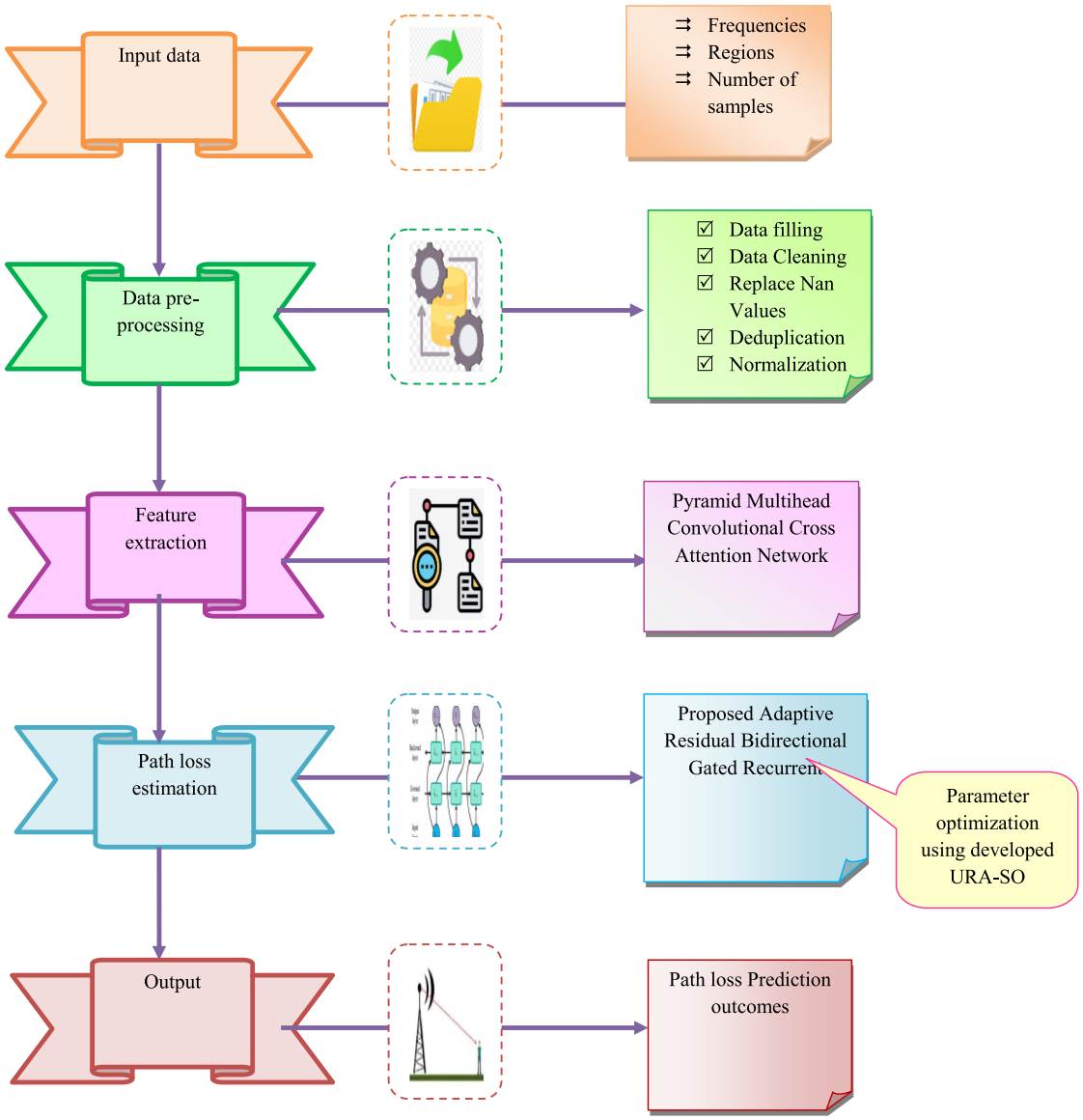

FIGURE 1 | Architectural framework of the proposed path loss estimation model in mm-Wave communication.

TABLE 2 | Region and sample details used for path loss estimation.

Regions	Number of samples
Port Harcourt	2800
Lagos Island	3050
Ibadan	3100
Abuja	2800
Total	11,750
Training and testing split	
Description	**Count**
Total samples	11,750
Training samples (75%)	8813
Testing samples (25%)	2937

specified as w. The dataset details are listed in Table 2 and in Figure 2.

4 | Improvement in Estimation Performance Through Preprocessing and Feature Extraction Mechanisms

4.1 | Data Preprocessing

The preprocessing techniques greatly improve the dataset's quality and usability for accomplishing predictions as well as classification. The input data C_w is transformed into a clean and consistent format by removing redundancy, filling in gaps, and correcting inconsistencies. Normalization also enhances the input characteristics for better generalization. Overall, this pre-processing procedure helps to lower the noise and bias issues for performing stable path loss estimation.

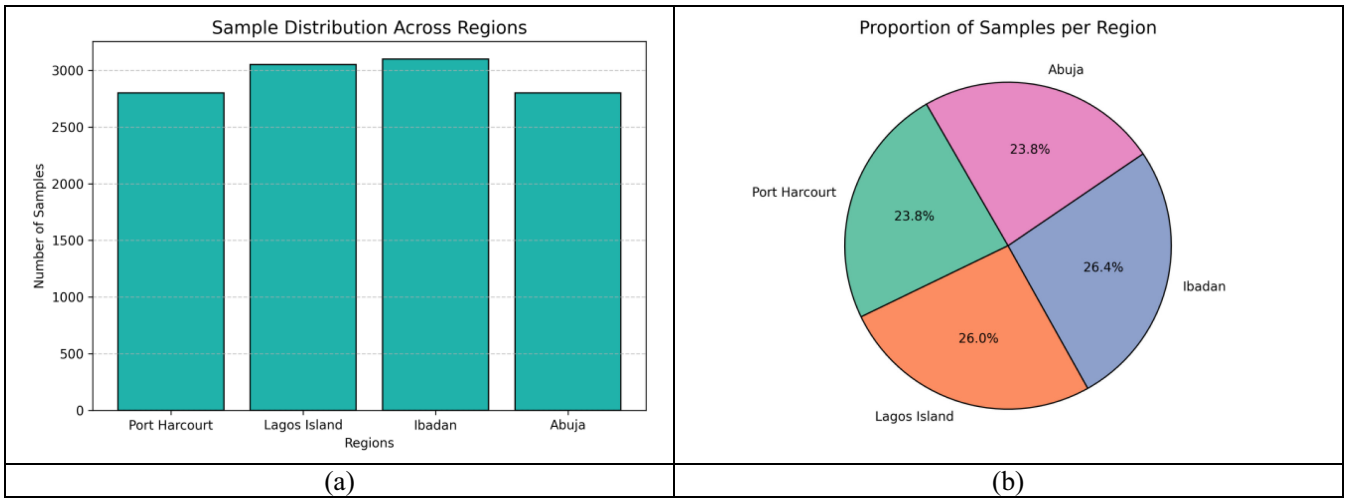

FIGURE 2 | Region and sample details of the suggested path loss estimation model.

4.1.1 | Data Filling

Adding missing or insufficient entries to the dataset by statistical or logical methods is known as data filling. Missing values for one or more of the samples' attributes may be present in the raw dataset. This missing value problem includes network errors or errors from the device that is used to collect the data. Eliminating the sample with the missing value is one method of resolving the missing value issue. To fill in the missing data, it is beneficial to use mean and median computation to replace the missing values instead of removing them.

4.1.2 | Data Cleaning

Data cleansing is one of the preprocessing steps used to handle inaccurate and incomplete data. In order to reduce redundancy, the data cleaning procedure identifies unnecessary data and removes it. Noise issues in datasets are filtered during the data-cleaning process.

4.1.3 | Replace NaN Values

In this phase, the not a number (NaN) values are considered as undefined dataset elements. To prevent bias during the path loss prediction process, the NaN values are substituted with suitable values that take into account the mean or median.

4.1.4 | Deduplication

In order to maintain the integrity of collected data, deduplication is performed. This step ensures that all the duplicate entries are removed from the dataset. The computational overhead issues are rectified by removing the duplicate data.

4.1.5 | Normalization

Data are transformed into a consistent format via the normalization process. In this step, the numerical values are scaled to a common range like 0–1. Through this normalization process, every valuable feature in this data is equally beneficial to the learning process. The statistical properties of data are preserved even after normalization. The preprocessed data are indicated as N_v^{PRE}.

4.2 | Deep Feature Extraction Step

The deep feature extraction process is beneficial for identifying both local and global variations in high-dimensional data. In this work, the PMC-CANet model is used for feature extraction, as it gives importance to the most pertinent characteristics across spatial and contextual dimensions by combining attention-based techniques with pyramid convolution procedures. The preprocessed data are considered as input N_v^{PRE} for this procedure.

The pyramid multihead convolutional architecture [29] employed a sequence of three convolutional layers with kernel dimensions that are set to 17, 11, and 7, whereas the output channel counts were 128, 256, and 512. ReLU activation and a dropout layer were used to form a convolutional block after each of these convolutional operations. After every convolution block, a multihead self-attention module with eight heads was incorporated. The mathematical formulation for this attention mechanism is provided in Equations (1) and (2).

$$MH(Y,P,D) = C\left(H_1,.\ \ldots H_N\right)E^G \tag{1}$$

$$H_a = Q\left(Y_a,P_a,D^a\right) = \text{Soft max}\left(\frac{Y_a P_a^T}{\sqrt{b_j}}\right)D_a \tag{2}$$

In the above equation, query Y, key P, and value D are represented. Next, E^G denotes the extracted global features. The concatenate function is given as $C()$. Head 1 is signified as H_1, and a^{th} head is denoted as H_a. The term Q is attention. To create a single output feature, a fully connected layer was employed in the final stage of the model. Its input was derived by aggregating the outputs from various scales, and each of them is modulated by a trainable weight parameter.

In the cross-attention model [30], two separate networks apply attention to one another, which allows for the extraction of more informative feature representations. Initially, the input is passed through both networks to obtain two sets of feature maps. Negative activations are eliminated by employing ReLU activation layers to process these maps. Following activation, the feature sets are passed through a transition layer to align their shapes to ensure compatibility for further operations. Instead of computing attention via the conventional outer-product, this approach applies an element-wise Hadamard product to derive the cross-attention feature maps. By highlighting useful features and minimizing noise, this operation allows the model to focus on regions with higher pathogenic information. The resulting cross-attention feature maps are then concatenated with the original outputs from both networks to form the final representation. The cross-attention process is mathematically derived in Equations (3) and (4).

$$R_{Con} = C\left(R_X, R_U, R_V\right) \tag{3}$$

$$R_X = R_U \otimes R_V \tag{4}$$

Here, the cross-attention feature map is mentioned as R_X. The feature map of the network U and V is stated as R_U and R_V, respectively. The processed output is expressed as shown in Equation (5)

$$\delta = \mu N^i + N \tag{5}$$

Here, the output after the attention layer is termed as N^i, and the input given to the flatten layer is mentioned as δ. The multihead cross attention is defined in Equation (6).

$$\begin{aligned} MH\left(Y^f, P^g, D^h\right) &= C\left(H_1, \ldots H_N\right) L^e H_i \\ &= CA\left(Y^f L_i^f, P^g L_i^g, D^h L_i^h\right) \end{aligned} \tag{6}$$

In the above equation, the cross attention is specified as CA. The extracted features from PMC-CANet are mentioned as G_P^{fet}. A schematic view of PMC-CANet-assisted feature extraction is shown in Figure 3.

The PMC-CANet-based feature extraction model starts with input preprocessed data N, which goes through initial hierarchical feature extraction by a sequence of convolutional layers Conv 1×1 and Conv 3×3 to extract multiscale patterns. Later, these features are refined and given to the dual-stream attention modules, where Softmax and Matrix Multiplication (MatMul)

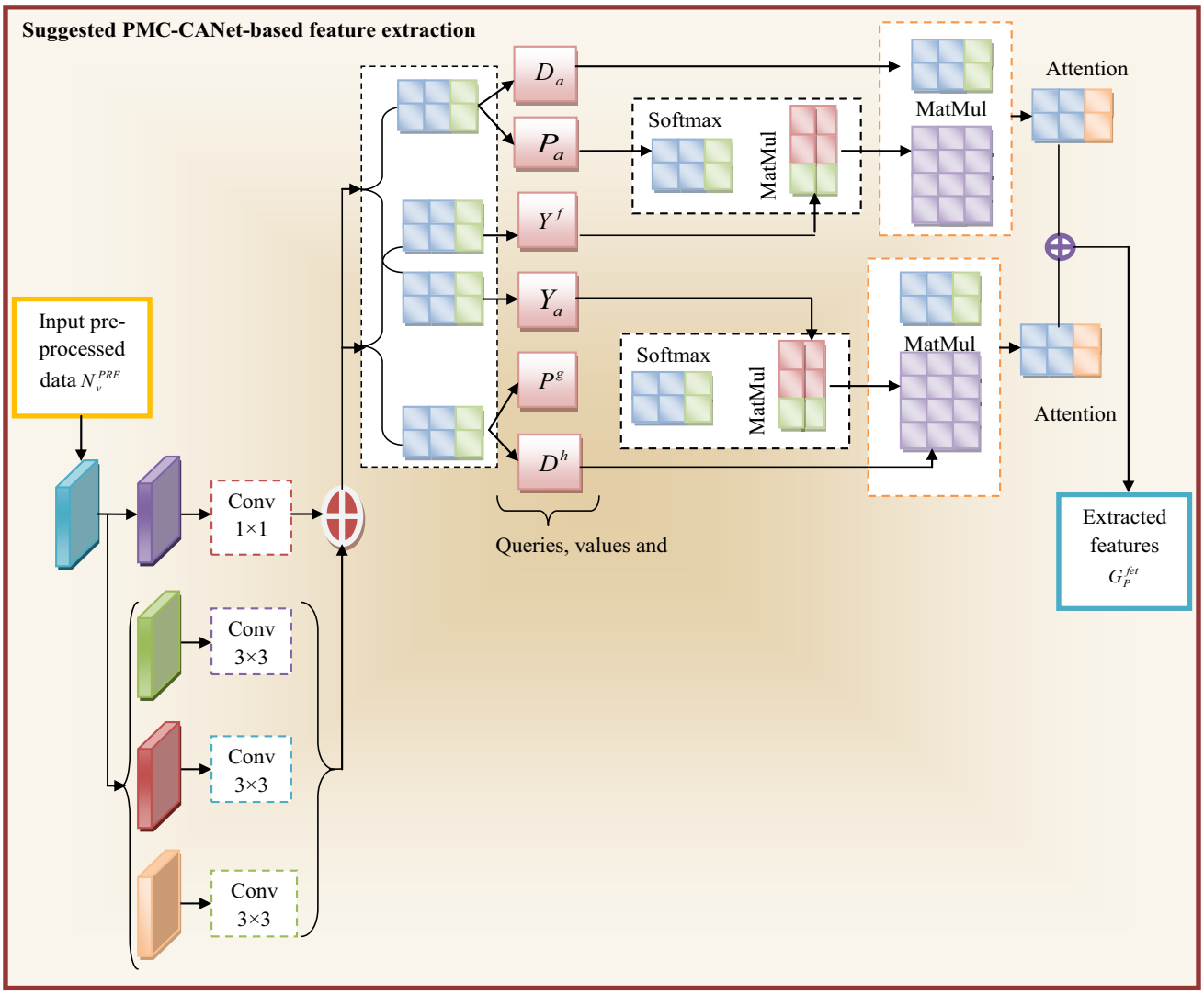

FIGURE 3 | Schematic view of PMC-CANet-assisted feature extraction.

operations calculate attention weights, enabling the model to carefully concentrate on suitable characteristics by considering those as queries and values, which improves their background significance. The processed outputs from these attention blocks are then combined, resulting in the extracted features, which symbolize discriminative, multiscale, and contextually-aware representations modified for downstream tasks.

4.3 | Developed URA-SO

The proposed AR-BiGRU model consists of several hyperparameters, and fine-tuning those helps to prevent overfitting issues. The purpose of this work is to introduce a URA-SO algorithm that optimizes model performance and increases its generalization capability. The URA-SO is used to efficiently fine-tune the parameters in a specified interval to reduce the inconsistency of estimating path loss. When parameters are properly optimized, the model effectively handles complicated conditions that cause path loss during the mmWave communication including the temporal as well as environmental factors. The optimization procedure leads to more accurate forecasts by changing the parameters in the AR-BiGRU network. The complexity issues are handled precisely by the optimal solution, as it manages issues like atmospheric conditions during the mm-Wave transmissions for determining the path loss. In SOA [31], the mathematical modeling of changes in sculpting materials is useful to improve the exploration phase by avoiding local optima and handling multimodal as well as hybrid optimization problems. Due to its inefficiency in the exploitation phase, this algorithm's convergence is slow, as it takes more time to find the solution. Therefore, the optimization efficiency of SOA is enhanced by updating the random parameter h as shown in Equation (7).

$$h = \frac{S2}{S1} \tag{7}$$

In the above derivation, the term S is the fitness function. Initially, the fitness values are sorted in ascending order based on the solutions. Then, the random integer is computed based on the first fitness $S1$ and last fitness $S2$ in the sorted list. Wider exploration, as well as exploitation, is possible by upgrading the random parameter based on the fitness computation. By sorting fitness functions in ascending order, the random number update process is not time-consuming, making it possible to achieve the best global solution efficiently without taking more time. The transition between the two phases in conventional SOA is improved by the proposed URA-SO. The algorithm's search efficiency is improved by the random number update, which prevents premature convergence. Algorithm 1 provides the pseudocode of the URA-SO algorithm.

5 | Description of Developed Deep Learning Mechanism for Path Loss Estimation With Its Objective Function

5.1 | Residual Bidirectional Gated Recurrent Unit

An R-BiGRU model is considered for estimating the path loss in mm-Wave wireless communication.

In the R-BiGRU network [32], the residual architecture utilizes skip connections, and it consists of two or three consecutive layers that include nonlinear functions such as batch normalization and ReLU. Preventing issues such as gradient vanishing and accuracy saturation problems can be achieved by using skip connections in the residual BiGRU. The output R attained from the residual module is given in Equation (8).

$$R = \zeta_u\left(g + V\left(Conv\left(\zeta_u(V(Conv(g)))\right)\right)\right) = \mathrm{Re}(g) \tag{8}$$

The ReLU nonlinearity is termed as ζ_u. The variables $\mathrm{Re}(m)$ and g indicate the residual block and the input of the module, respectively.

The standard GRU is designed to handle sequential or time-dependent data via the update gate and the reset gate. The update gate of the GRU is in charge of both the forget gate and the input gates of the LSTM network. However, standard GRUs function in only one direction, so they cannot process past and future information. The Bi-GRU architecture has been created to address this and does not necessitate separate memory cells.

Algorithm 1: Implemented URA-SO
Input: AR-BiGRU parameters
Assign the population
Assign the iteration M
Generate the populations
While until condition fulfilled
For $i = 1 to C$
Identify the pattern of sculpture
Estimate position of i^{th} sculptor
Evaluate objective function of current position
Update the position
Upgrade the random value h using Eq. (7)
Estimate the novel locations of i^{th} sculptor
Evaluate novel position objective function
Update the new locations
End For
End while
Attain the good outcomes
Output: Optimal steps per epoch in BiGRU SP_H^{BiGRU}, learning rate ED_P^{BiGRU} and hidden neuron count YT_N^{BiGRU}

The input sequence is simultaneously processed by a forward GRU and a backward GRU.

The operation of the update gate a_d is determined using Equation (9).

$$a_d = \sigma\left(X_a p_d + Y_a w_{d-1}\right) \tag{9}$$

Here, the hidden state and input are referred by the terms w_{d-1} and p_d, concurrently. The weight multiplied with the hidden state is termed as Y_a, and the weight multiplied by the input is specified as X_a. The forget gate is responsible for removing past information, which is mathematically expressed in Equation (10).

$$b_d = \sigma\left(X_b p_d + Y_b w_{d-1}\right) \tag{10}$$

The computation of the candidate hidden state is provided in Equation (11).

$$\hat{w}_d = \tanh\left(X_w p_d + b_d \otimes Y_w w_{d-1}\right) \tag{11}$$

Here, the weights are multiplied, and an element-wise multiplication $\otimes$ is done between the reset gate as well as the weight of the hidden state. These results are summed and sent to the nonlinear activation function tanh. The current hidden state is processed based on Equation (12).

$$w_d = (1\text{-}a_d) \otimes \hat{w}_d + a_d \otimes w_{d\text{-}1} \tag{12}$$

The hidden state in Bi-GRU is computed using Equation (13).

$$w_d = Concatenate\left(\overrightarrow{w_d}, \overleftarrow{w_d}\right) \tag{13}$$

At each time step d, the hidden state is computed by concatenating the hidden vectors $\overrightarrow{w_d}$ and $\overleftarrow{w_d}$, generated by the forward and backward GRUs, respectively. The delineated representation of the R-BiGRU model is showcased in Figure 4.

5.2 | Path Loss Estimation Using AR-BiGRU

The retrieved deep features G_P^{fet} are sent to the developed AR-BiGRU for estimating the path loss. The AR-BiGRU is capable of capturing temporal dependencies in forward and backward directions, which ensures precise handling of dynamic environmental changes. Vanishing gradient issues are rectified while enhancing information flow by the incorporation of residual connections. The intricate as well as nonlinear relations in the input features are precisely modeled by the suggested techniques. Bidirectional processing of the AR-BiGRU approach allows it to learn contextual dependencies from both historical and new patterns, which increases the accuracy of temporal sequence modeling. The residual learning and bidirectional flow of the proposed methods are useful in handling the features obtained from data that are taken in high-interference and urban regions. The BiGRU model can accurately learn spatial

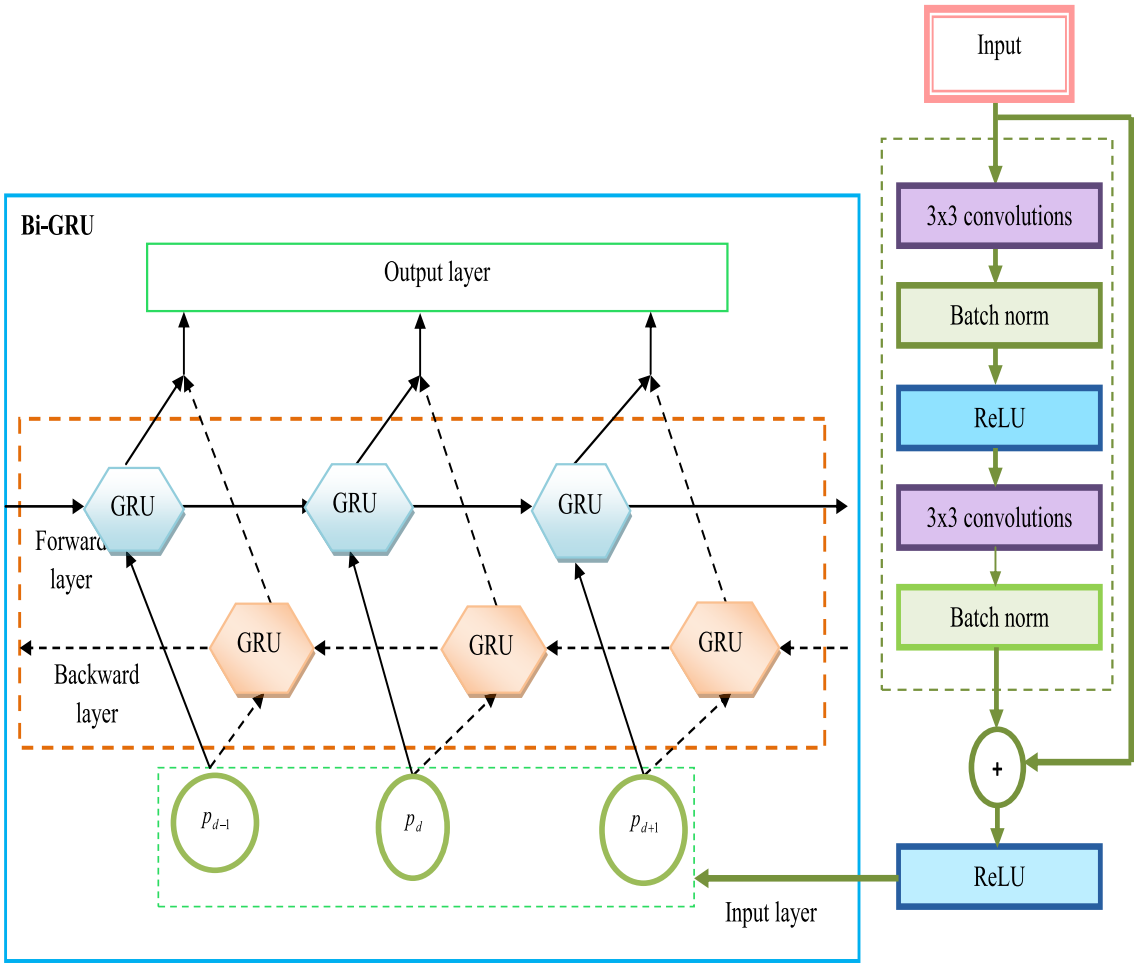

FIGURE 4 | Delineated representation of the R-BiGRU model.

patterns in path loss because the input data include location details. By analyzing the cause of path loss estimation, communication via mm-Wave wireless transmission networks is enhanced. The path loss PL is estimated based on Equation (14).

$$\overline{PL(c)}(dB) = \beta + \overline{\gamma}.10\log_{10}(c) \tag{14}$$

Here, the term c indicates the distance. The linear slope and floating intercept are defined by the variables $\overline{\gamma}$ and β; they are computed using Equations (15) and (16).

$$\overline{\gamma} = \frac{\sum_j^z \left(c_j - \overline{c}\right) \times \left(PL_j - \overline{PL}\right)}{\sum_j^z \left(c_j - \overline{c}\right)^2} \tag{15}$$

$$\beta(dB) = \overline{PL}(dB) - \overline{\gamma}.\overline{10\log_{10}(c)} \tag{16}$$

In the above notations, the average path loss and the average distance is mentioned as $\overline{PL}$ and $\overline{c}$, respectively. The dependence of frequency is estimated using Equation (17).

$$\overline{PL(c,q)}(dB) = \beta + \overline{\gamma}.10\log_{10}(c) + \alpha.20\log_{10}(c)\left(\frac{q}{q_f}\right) \tag{17}$$

Here, the center frequency is given as q_f. Target frequency and the frequency-dependency factor are signified as q and α, respectively. The ratio of frequency deviation is given as $\frac{q}{q_f}$. Visualization of the proposed URA-SO-AR-BiGRU-based path loss estimation in mm-Wave wireless communication is shown in Figure 5.

5.3 | Objective Function

The URA-SO algorithm is used to tune the important hyperparameters of the BiGRU to improve estimation accuracy. Optimizing parameters like hidden neurons and steps per epoch can improve the efficiency of the AR-BiGRU model. These parameters are essential for managing the model's efficiency to

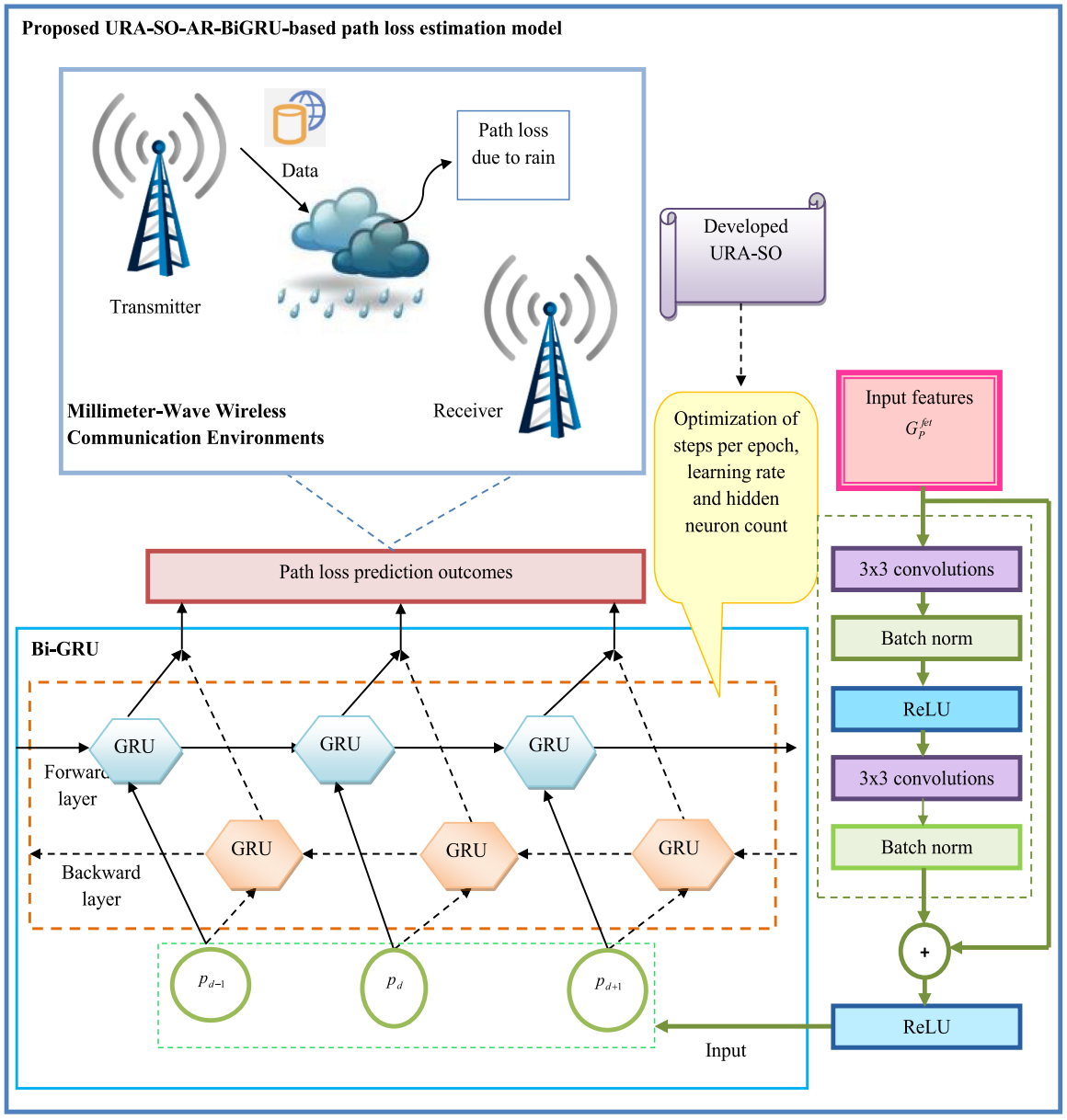

FIGURE 5 | Visualization of proposed URA-SO-AR-BiGRU-based path loss estimation in mm-Wave wireless communication.

study more details from the extracted features while overcoming the overfitting issues. The optimization process improves the capacity of AR-BiGRU in learning temporal and spatial features to estimate path loss. Similarly, the complexity of the training process is also managed by the optimization of the learning rate and steps per epoch.

The objective function W_n of URA-SO-AR-BiGRU-based path loss estimation is provided in Equation (18).

$$W_n = \underset{\{SP_H^{BiGRU}, ED_P^{BiGRU}, YT_N^{BiGRU}\}}{\arg\min} \left(\frac{1}{Accuracy} + MSE \right) \tag{18}$$

In the above equation, the steps per epoch in BiGRU are tuned in between $[100 - 500]$, and it is specified as SP_H^{BiGRU}. The optimized learning rate and hidden neuron count in BiGRU are referred as ED_P^{BiGRU} and YT_N^{BiGRU}; both lies in the interval of $[0.01 - 0.99]$ and $[5 - 255]$.

- The mathematical computation for accuracy is given in Equation (19).

$$Accuracy = \frac{G^{TP} + G^{TN}}{G^{TP} + G^{TN} + F^{FP} + F^{FN}} \tag{19}$$

- Mean squared error (MSE) is provided in Equation (20).

$$MSE = \frac{1}{J} \sum_{z=1}^{j} \left(H_z - \hat{H}_z \right)^2 \tag{20}$$

Here, the total data point is given as J. The true negative, true positive, false positive, and false negative scores are specified as G^{TN}, G^{TP}, F^{FP}, and F^{FN}, correspondingly. The predicted and actual value is stated as $\hat{H}_z$ and H_z.

6 | Performance Evaluation and Discussions

6.1 | Experimental Setup

The path loss technique implementation was carried out using Python. Factors like maximum iteration, chromosome length, and number of population were assigned as 50, 3, and 10 during the execution. Existing models were compared to validate the efficiency of the path loss determination process. Heuristic approaches including Northern Goshawk Optimization (NGO) [33], Carpet Weaver Optimization (CWO) [34], Remora Optimization Algorithm (ROA) [35], and SOA [31] were analyzed. Also, state-of-the-art models like DNN [19], VGG-16 [20], RNN [24], and R-BiGRU [32] were considered for verifying the accuracy of the proposed approach. The feature extraction performance was compared with CNN [20], Autoencoder [36], PCA [37], and CAE [38].

6.2 | Performance Measures

Different performance indicators are used to compare the performance of path loss estimation against different techniques. The metrics used for validating the prediction performance are calculated using Equations (21–24).

$$MPE = \frac{\sum \frac{\hat{H}_z - H_z}{\hat{H}_z}}{J} \tag{21}$$

$$RMSE = \sqrt{\sum_{z=1}^{j} \frac{\left(\hat{H}_z - H_z \right)^2}{j}} \tag{22}$$

$$MASE = \frac{1}{j} \sum_{z=1}^{j} \left(\frac{\left| H_z - \hat{H}_z \right|}{\frac{1}{j-1} \sum_{z=2}^{j} \left| H_z - H_{z-1} \right|} \right) \tag{23}$$

$$SMAPE = \frac{1}{j} \sum_{z=1}^{j} \frac{\left| H_z - \hat{H}_z \right|}{\left(\frac{|H_z| + |\hat{H}_z|}{2} \right)} \tag{24}$$

6.3 | Performance Validation of Feature Extraction

The performance of PMC-CANet in feature extraction is compared among other feature retrieval mechanisms, and the results are displayed in Figure 6. The proposed PMC-CANet's relief score surpasses those of the existing methods due to its ability to capture both local and global contextual dependencies from the preprocessed data. The multihead cross-attention module improves the processing ability for finding the most important features. The pyramid convolution structure is useful for performing multiresolution analysis in the input data, which allows for the retrieval of relevant features necessary for accurate path loss estimation. The CAE are not capable of handling the spatial fluctuation issues present in mmWave data. At the 4th head convolution, the chi-squared static of the presented PMC-CANet in feature extraction is 32.81%, 25%, 18.75%, and 6.25% higher than CNN, Autoencoder, PCA, and CAE. The PCA model is a linear approach, and it might miss the relevant features that are useful for determining the path loss. Similarly, the existing approaches like the autoencoder are capable of learning nonlinear embeddings, but they often suffer from overfitting issues and miss the spatial information. Overfitting issues can be avoided even with high-dimensional inputs due to the pyramid convolution process that uses a multihead cross-attention mechanism.

6.4 | Performance Validation of Path Loss Estimation

The performance validation of the proposed URA-SO-AR-BiGRU for path loss estimation is conducted among prior techniques and algorithms as shown in Figures 7 and 8. The efficiency is tested by considering different predictive measures among activation functions. Nonlinear transformations in the retrieved features are precisely handled by the suggested AR-BiGRU mechanism, and it is confirmed by the high accuracy scores attained by the model. The RMSE of the URA-SO-AR-BiGRU in path loss estimation while analyzing with Tanh function is 13%, 25%, 44%, and 41% better than DNN, VGG-16, RNN, and R-BiGRU models. The AR-BiGRU model's adaptive temporal

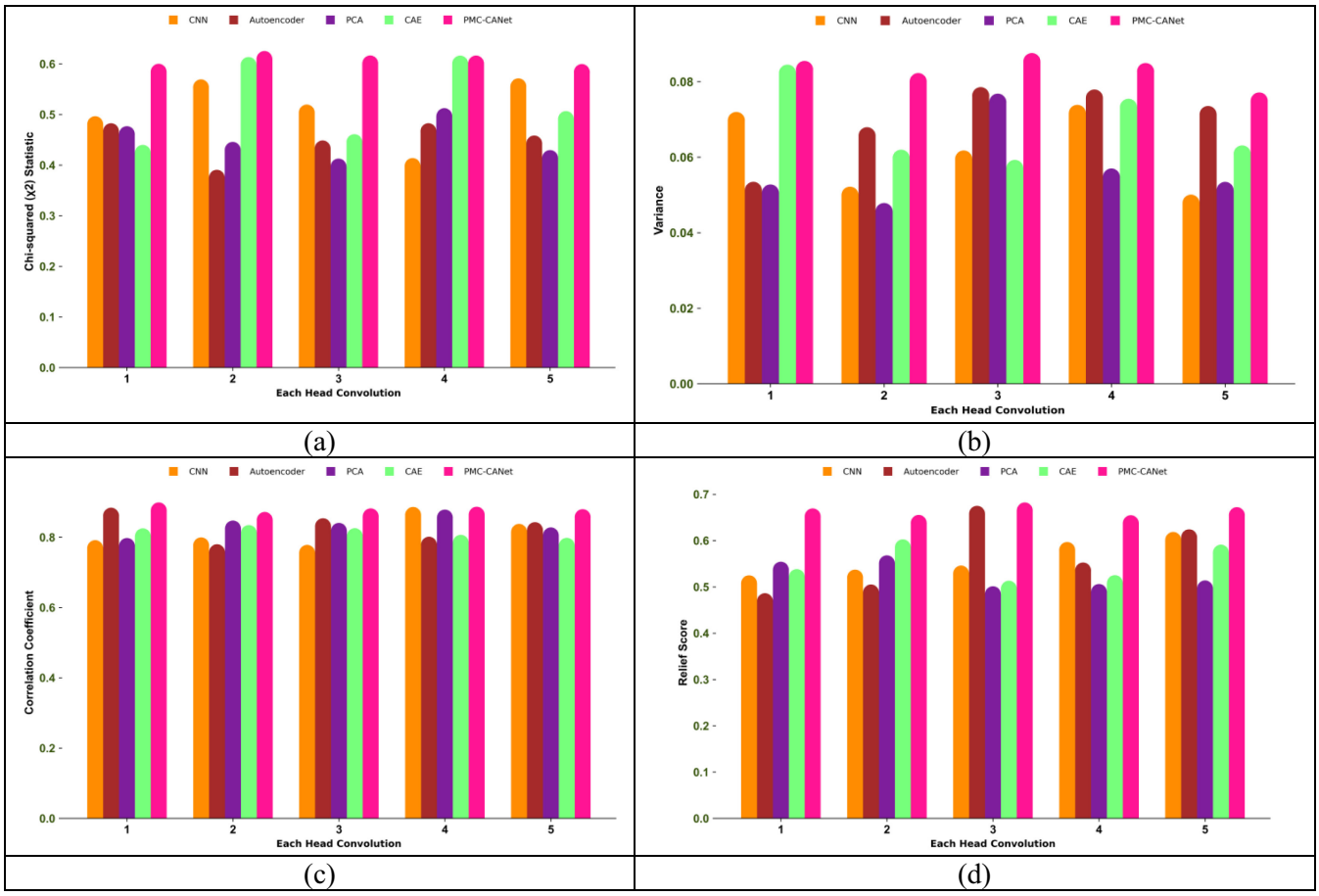

FIGURE 6 | Performance validation of feature extraction model with respect to (a) chi-squared statistic, (b) variance, (c) correlation coefficient, and (d) relief score.

modeling and hierarchical spatial characteristics make it possible to minimize prediction errors, as demonstrated by fewer MEP, SAMPE, and MASE outcomes. The inability to solve the vanishing gradient problems in existing models resulted in high prediction errors. Moreover, reducing computational complexity issues and obtaining accurate forecasting results were achieved by parameter tuning and an effective feature learning process before path loss estimation.

6.5 | Accuracy Computation of Developed URA-SO-AR-BiGRU

The accuracy computation of the developed URA-SO-AR-BiGRU among conventional algorithms and techniques is given in Table 3. In the epoch-wise performance comparison, the proposed URA-SO-AR-BiGRU attained an accuracy of 97.12%, whereas DNN, VGG16, RNN, and R-BiGRU obtained an accuracy of 82.4%, 92.56%, 87.12%, and 93.36% at an epoch count of 15. These results confirm the rapid learning ability of the suggested URA-SO-AR-BiGRU in precisely handling diverse path loss patterns. The adaptability of handling nonlinear data by the conventional models is low as per the accuracy analysis. The accuracy of RNN in estimating path loss is low due to vanishing gradient issues. The accuracy of URA-SO-AR-BiGRU is 94.24%, 96.8%, 97.12%, and 96.64% for the epoch range of 5, 10, 15, 20, and 25. Although the R-BiGRU model outperforms other approaches, URA-SO-aided optimization improves its performance in path loss estimation. Thus, it overcomes the bias and high-dimensional data handling issues more efficiently.

6.6 | Convergence Outcome

The convergence outcome of the proposed URA-SO-AR-BiGRU-based path loss estimation model is given in Figure 9. The convergence graph analysis is helpful to demonstrate the optimal efficiency of the URA-SO algorithm in fine-tuning the parameters of the BiGRU model. Performance degradation issues are rectified by the stable convergence achieved by the URA-SO-aided optimization. By using sorted fitness-based random number computation, the proposed URA-SO algorithm enhances the search process. This optimization process lowers the computational overhead issues while increasing the precision of path loss prediction. As per the results, conventional algorithms like NGO-R-BiGRU, ROA-R-BiGRU, CWO-R-BiGRU, and SOA-R-BiGRU exhibit delayed convergence among the iteration ranges. This optimization efficacy is beneficial in obtaining better path loss estimation outcomes by enhancing the bidirectional processing ability of the BiGRU model.

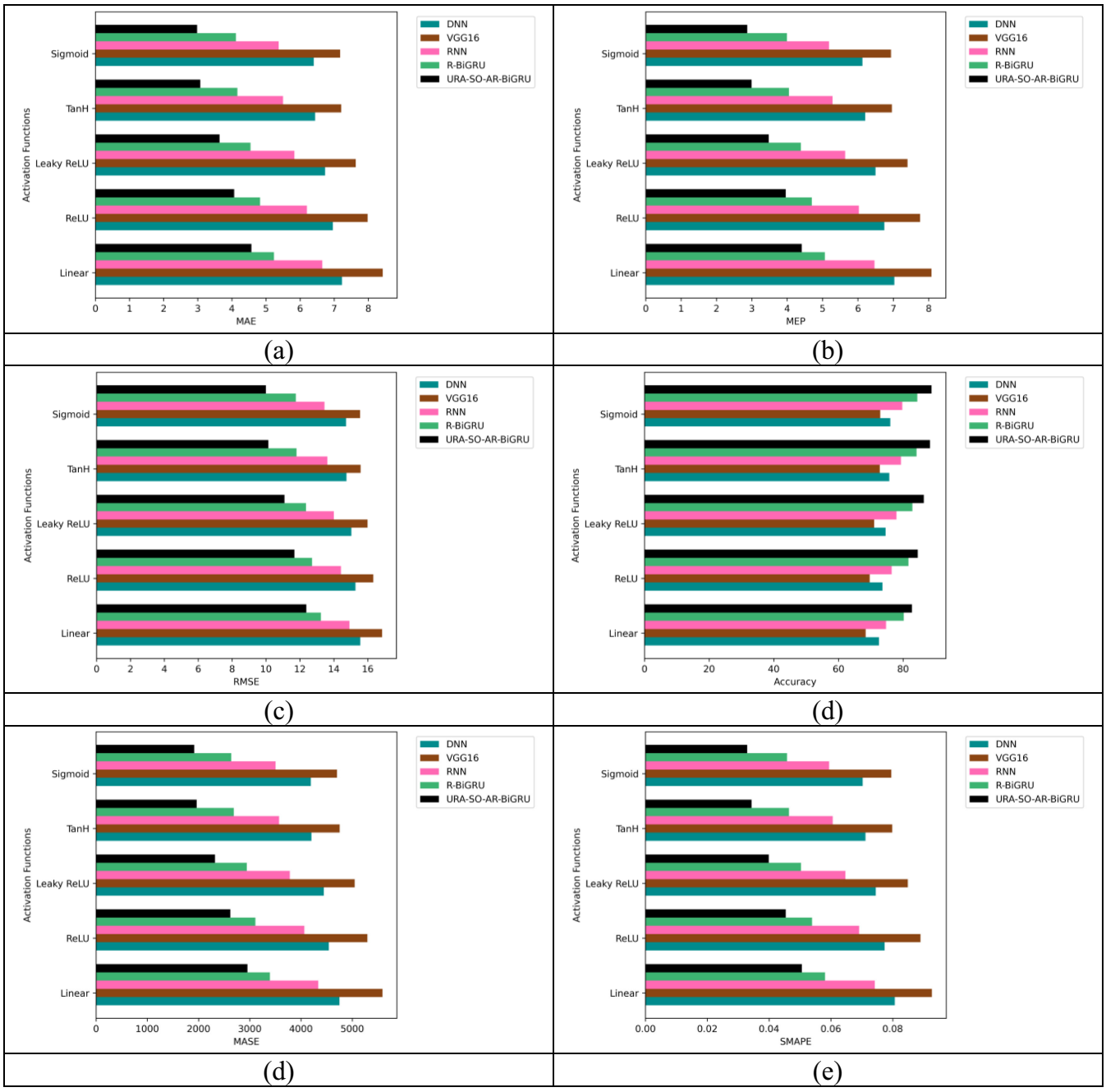

FIGURE 7 | Performance validation of introduced path loss estimation technique among existing methods in terms of (a) MAE, (b) MEP, (c) RMSE, (d) accuracy, (e) MASE, and (f) SMAPE.

6.7 | Statistical Results

The statistical evaluation of the introduced URA-SO-AR-BiGRU is conducted among other heuristic optimization algorithms, and the results are provided in Table 4. The best statistical measure of the presented approach is 0.791, whereas the mean and median are 0.9411 and 0.791, respectively. The URA-SO algorithm consistently achieves optimal performance by obtaining better mean outcomes than other algorithms. The standard deviation of just 0.4643 confirms the stable convergence behavior of the developed URA-SO over multiple iterations. Existing algorithms such as NGO NGO-R-BiGRU, ROA-R-BiGRU, CWO-R-BiGRU, and SOA-R-BiGRU often get trapped in local minima, as they result in less convergence reliability. The NGO and CWO algorithms reach the worst values of 5.19 and 4.52, which indicate the unstable search patterns of those algorithms in the problem space. Parameter tuning becomes less efficient in traditional SOA due to the absence of a fitness-based random integer update process.

6.8 | Statistical Testing on the Proposed Model's Accuracy Across Traditional Models

The statistical testing on proposed URA-SO-AR-BiGRU is tabulated in Table 5. This computation supports to establish the statistical enhancement of the developed URA-SO-AR-BiGRU over classical models. Here, the ANOVA *F*-statistic value of 11.4263

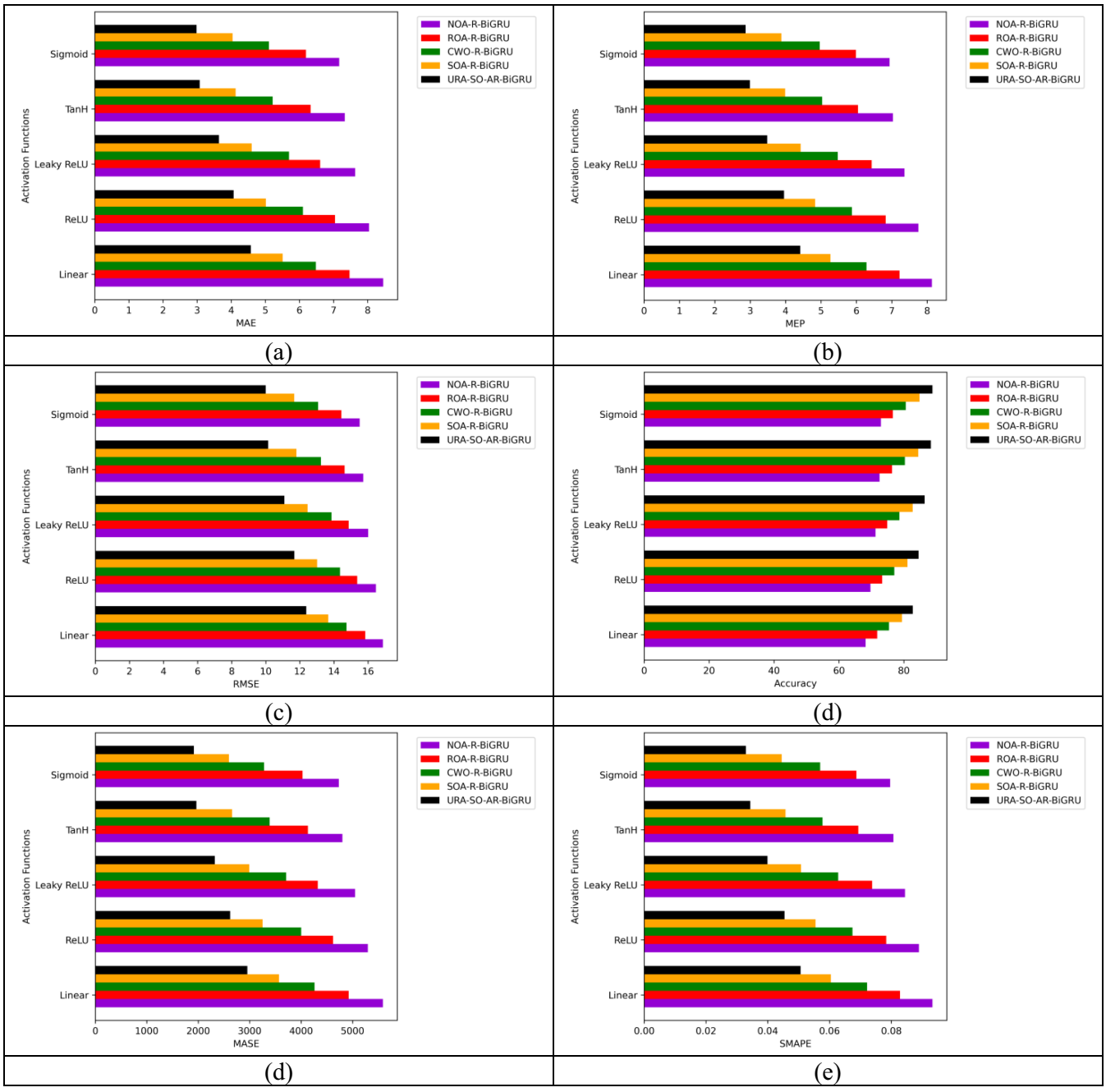

FIGURE 8 | Performance validation of the introduced path loss estimation model among conventional algorithms for (a) MAE, (b) MEP, (c) RMSE, (d) accuracy, (e) MASE, and (f) SMAPE.

shows that there is a major total difference in the mean accuracy of the frameworks being compared. Moreover, while comparing the proposed model versus VGG16, the t statistic of 3.76 and a p value of 0.0055 demonstrate that the developed URA-SO-AR-BiGRU accuracy is notably superior to the VGG16 model. Therefore, it is proven that the proposed URA-SO-AR-BiGRU performs better compared to other traditional approaches.

6.9 | Performance Comparison of the Proposed Mechanism Against Baselines

Figure 10 shows the performance comparison of the proposed URA-SO-AR-BiGRU compared to other baselines. This analysis is useful to estimate the efficiency and the performance of the suggested model. Here, the traditional SVR model attains the highest RMSE value of 0.5%, leading to a poorer fit to the data and less accuracy. However, the proposed URA-SO-AR-BiGRU model attains the lowest RMSE value of 0.1%, indicating more précised path loss estimation, as the estimations are much nearer to the actual path loss values. Hence, it is proven that the proposed model outperforms other baselines.

6.10 | Ablation Validation on Proposed URA-SO-AR-BiGRU

In Table 6, ablation observation of the developed URA-SO-AR-BiGRU is given. Generally, this computation is employed to assess the robustness. The table shows that the traditional

TABLE 3 | Accuracy computation of suggested path loss estimation model.

Accuracy					
Epoch	**NGO-R-BiGRU [33]**	**ROA-R-BiGRU [35]**	**CWO-R-BiGRU [34]**	**SOA-R-BiGRU [31]**	**URA-SO-AR-BiGRU**
Algorithm comparison					
5	83.04	84.96	84.48	87.52	94.24
10	84.08	86.88	87.04	91.52	96.8
15	84.16	83.12	84.32	84.96	97.12
20	83.04	91.04	89.68	88.56	96.64
25	92	92.32	80.96	90.72	94.48
Classifier comparison					
Epoch	**DNN [19]**	**VGG16 [20]**	**RNN [24]**	**R-BiGRU [26]**	**URA-SO-AR-BiGRU**
5	84	86.48	86.8	88.96	94.24
10	84.32	93.04	83.6	88.32	96.8
15	82.4	92.56	87.12	93.36	97.12
20	78.72	84.56	83.2	88.88	96.64
25	89.36	89.36	91.12	89.2	94.48

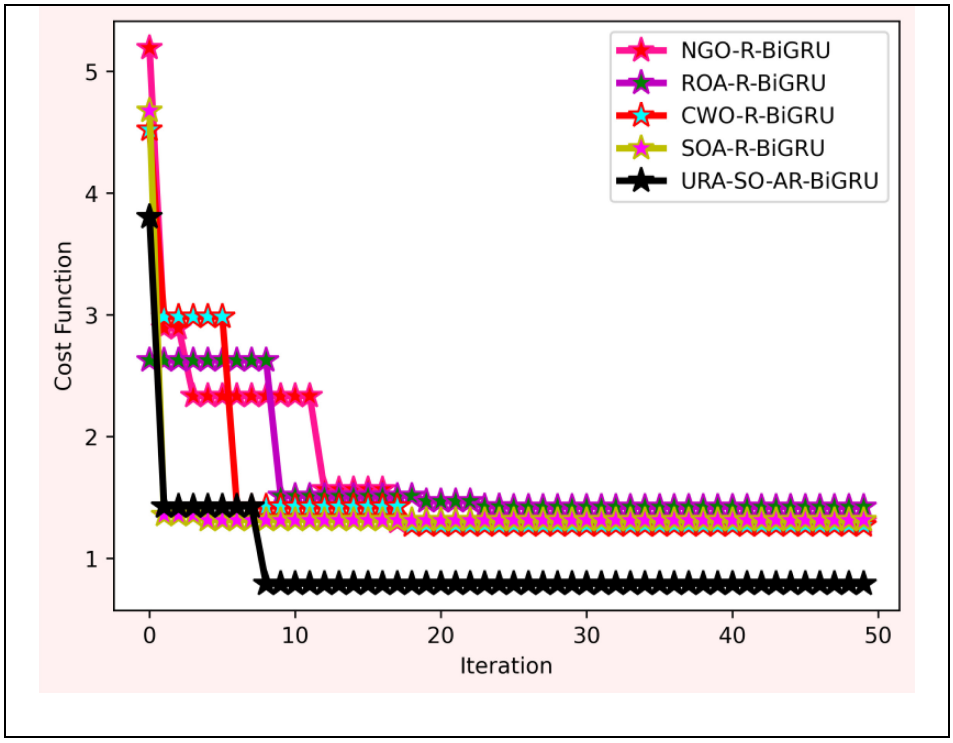

FIGURE 9 | Convergence outcome of suggested path loss estimation model.

TABLE 4 | Statistical results of suggested path loss estimation model.

Statistical measures	NGO-R-BiGRU [33]	ROA-R-BiGRU [35]	CWO-R-BiGRU [34]	SOA-R-BiGRU [31]	URA-SO-AR-BiGRU
Best	1.300417	1.424001	1.271852	1.315795	0.791963
Standard deviation	0.691655	0.453494	0.660147	0.470547	0.464341
Mean	1.654694	1.662579	1.545479	1.385501	0.941122
Worst	5.194203	2.627655	4.523694	4.678629	3.803367
Median	1.300417	1.426231	1.271852	1.315795	0.791963

PMC-CANet-Res-BiGRU model attains 83.92% of performance, which is lower than other models, indicating minimized data rates, improved energy usage, and poor device localization. However, the developed URA-SO-AR-BiGRU attains better accuracy as 97.12%, which helps to accomplish better outcomes without errors. Analysis displayed that the developed URA-SO-AR-BiGRU gained more precise path loss estimation than prior schemes.

6.11 | State of the Art Comparison of the Proposed Scheme

Table 7 establishes the state-of-the-art comparison of the implemented mechanism. This observation supports demonstrating the performance of each mechanism, providing a greater understanding. The traditional Unet [27] model attains 85.5% of accuracy, which leads to poor decision-making and ineffective performance. However, the developed URA-SO-AR-BiGRU model achieves 96.64% accuracy, indicating reliable decision-making and enhanced efficiency. Thus, the suggested mechanisms accomplish greater path loss than the baseline schemes.

TABLE 5 | Statistical validation of the proposed model's accuracy across traditional models.

ANOVA *F* statistic,	*p* value: F_onewayResult(statistic = 11.42632052178628)
T-test URA-SO-AR-BiGRU versus DNN	$t = 6.63$, $p = 0.0002$
T-test URA-SO-AR-BiGRU versus VGG16	$t = 3.76$, $p = 0.0055$
T-test URA-SO-AR-BiGRU versus RNN	$t = 6.08$, $p = 0.0003$

6.12 | Discussion on Proposed Model Interpretability and Deployment Feasibility

Table 8 provides the deployment analysis of the developed URA-SO-AR-BiGRU. The table shows that the model has a total size of 512.3 KB. And the proposed model took less training time of 2.134 s, indicating that the suggested URA-SO-AR-BiGRU framework is comparatively lightweight and resourceful. Moreover, a decreased model size and rapid inference time specify superior potential for deployment in a range of scenarios. Also, less training time leads to lower computational cost and enhanced accessibility. Moreover, an analysis on feature importance for path loss prediction is provided in Figure 11. This analysis shows the relative significance of diverse features in path loss prediction. Feature 3 attains a greater importance value of more than 0.4%, and Feature 2 attains an importance value of 0.32%. The analysis of model size, inference time, and feature importance emphasize the critical

TABLE 6 | Ablation observation on developed scheme.

Model/metrics	Accuracy
PMC-CANet	90.64
PMC-CANet -Res-BiGRU	83.92
PMC-CANet -AR-BiGRU	86.4
PMC-CANet - URA-SO-AR-BiGRU	97.12

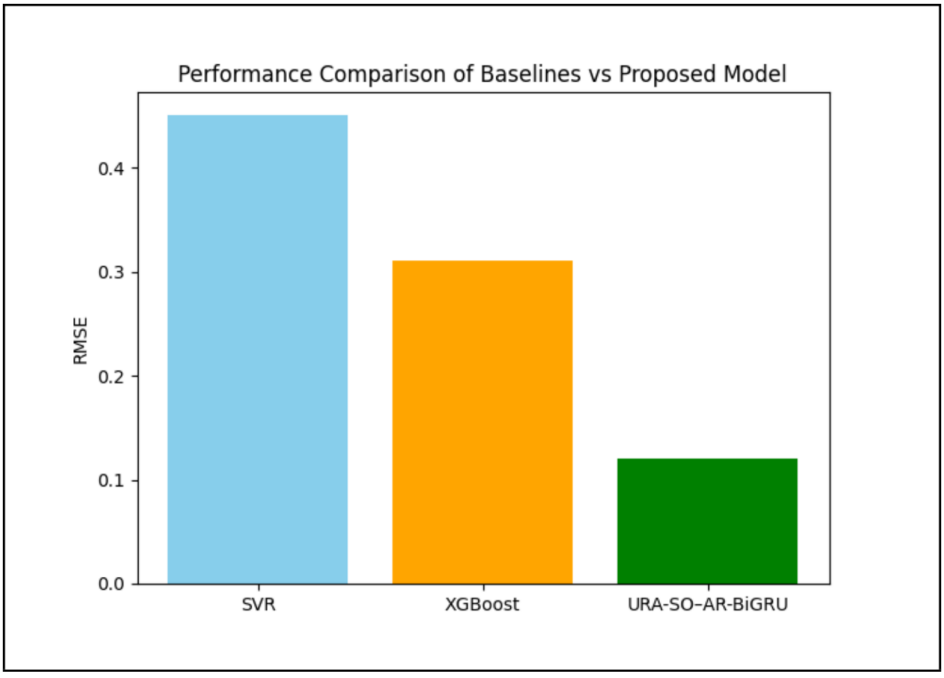

FIGURE 10 | Performance comparison of the proposed URA-SO-AR-BiGRU against baselines.

practical considerations for deploying models, emphasizing the necessitate to balance model intricacy and computational efficiency, performance, and interpretability to guarantee effective real-world application.

6.13 | Analysis of the Proposed Scheme Through Regression Performance Metrics

Figure 12 provides the performance analysis of the proposed model using four common regression error metrics, namely, MAE, RMSE, MAPE, and R squared (R^2). The figure shows that the proposed model attains a 1.37% MAE error value, which shows the absolute difference between the forecasted and the actual value, leading to effective and accurate outcomes. Hence, the validation results state that the developed URA-SO-AR-BiGRU is better at estimating the path loss than the prior mechanism.

TABLE 7 | State-of-the-art comparison of the proposed models.

Epoch	DNN [26]	Unet [27]	RF-SVR [28]	URA-SO-AR-BiGRU
5	91.92	85.92	90	94.24
10	90.72	84.88	86.96	96.8
15	92.56	91.84	86.88	97.12
20	85.52	85.84	92.32	96.64
25	88.16	89.52	90.88	94.48

TABLE 8 | Deployment analysis of the proposed models.

Metric	Value
Model size (KB)	512.3 KB
Training time (s)	2.134 s
Inference time (ms/sample)	0.25 ms

6.14 | Discussion

6.14.1 | Generalizability and Robustness of URA-SO-AR-BiGRU

The proposed URA-SO-AR-BiGRU has the capability to generalize well among varied wireless communication environments as maintaining robustness under varying operational conditions, in which the PMC-CANet feature extractor significantly contributes to generalizability. Its pyramid convolutional formation extracts multiscale differences, whereas the multihead cross-attention mechanism highlights the suitable contextual aspects. Therefore, the system learns local as well as global signal features, which could transfer efficiently to diverse settings, including cities, exurban, and domestic scenarios. Moreover, the AR-BiGRU with residual connections enhances robustness to chronological dependency in sequential and spatial data. By capturing mutual information exchange and ensuring smooth gradients, the model has the ability to adjust to rapid environmental variability, like atmospheric attenuation, which is frequent in practical mmWave deployments. Furthermore, the URA-SO strategy strengthens the robustness of the proposed URA-SO-AR-BiGRU by effectively tuning the hyperparameters of the AR-BiGRU model. This adaptive fine-tuning averts underfitting in uncomplicated settings and overfitting in intricate

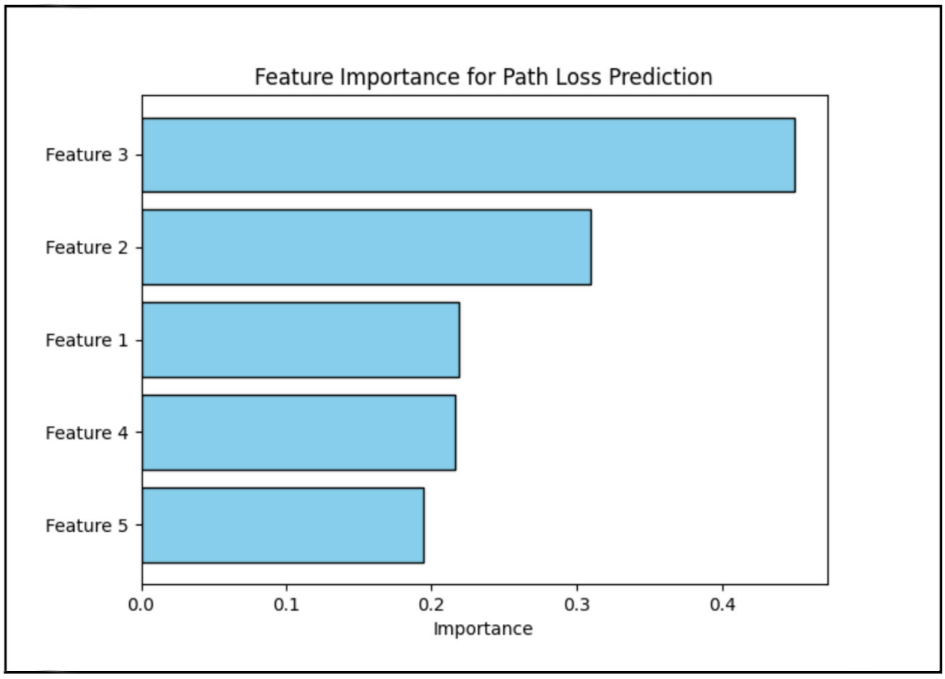

FIGURE 11 | Analysis of feature importance for path loss prediction.

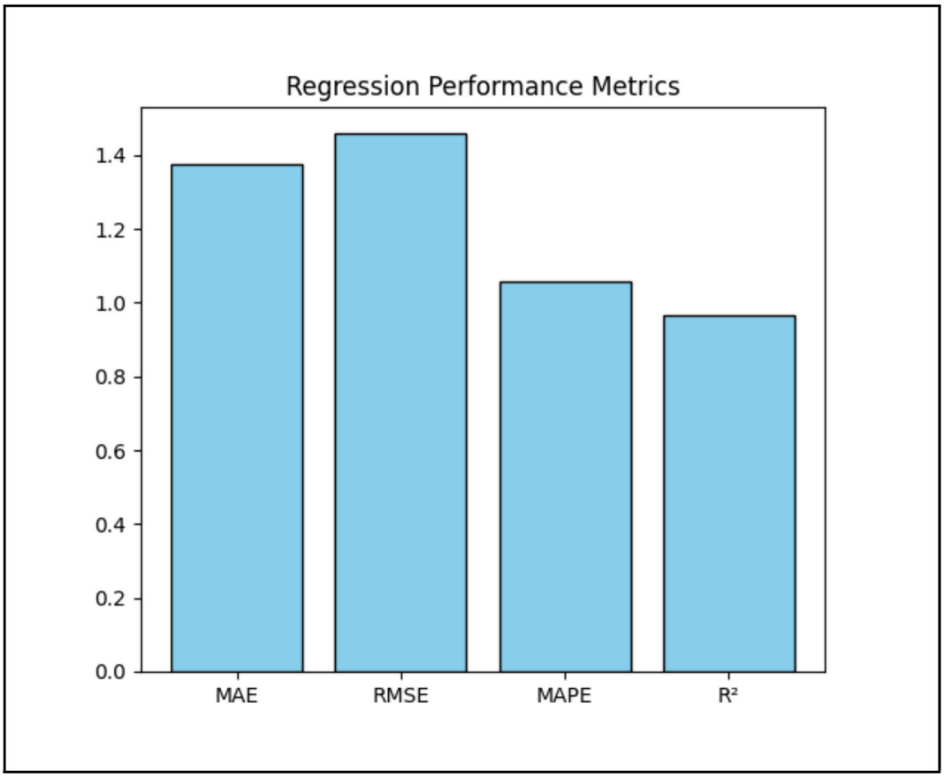

FIGURE 12 | Analysis on regression performance measures in the proposed URA-SO-AR-BiGRU.

settings, guaranteeing reliable performance across datasets with a varied degree of variations. Overall, the developed model maintains equilibrium on accuracy, scalability, and resilience, which makes it compatible for exploitation in 5G networks.

7 | Conclusion

An advanced deep learning technique was developed for accurate path loss estimation in mm-Wave wireless communication environments. After data collection, preprocessing was performed to clean and normalize the data for improving the prediction quality. The refined data were then passed through the PMC-CANet model for extracting features. To extract relevant features at multiple scales, this model utilized a multihead cross-attention mechanism and a pyramid convolution module. These deep features were passed to AR-BiGRU to estimate path loss. The URA-SO algorithm was employed to optimize the AR-BiGRU parameters for better performance. The entire performance of the suggested model was evaluated through experimental analysis against existing methods. The accuracy of the designed URA-SO-AR-BiGRU is 97.12% at an epoch range of 15. The resultant outcomes confirm the adaptability and robustness of the implemented URA-SO-AR-BiGRU in path loss estimation of mm-Wave wireless communication. Ensemble models will be considered in the future to enhance the accuracy of path loss estimation. The root cause of the path loss will be analyzed to minimize the path loss issues with advanced deep machine learning approaches.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article, as no datasets were generated or analyzed during the current study.

References

1. M. K. Elmezughi, O. Salih, T. J. Afullo, and K. J. Duffy, "Path Loss Modeling Based on Neural Networks and Ensemble Method for Future Wireless Networks," *Heliyon* 9, no. 9 (2023): e19685.

2. S. K. Hinga and A. A. Atayero, "Deterministic 5G mmWave Large-Scale 3D Path Loss Model for Lagos Island, Nigeria," *IEEE Access* 9 (2021): 134270–134288.

3. C. Nguyen and A. A. Cheema, "A Deep Neural Network-Based Multi-Frequency Path Loss Prediction Model From 0.8 GHz to 70 GHz," *Sensors* 21, no. 15 (2021): 5100.

4. M. Bal, A. Marey, H. F. Ates, T. Baykas, and B. K. Gunturk, "Regression of Large-Scale Path Loss Parameters Using Deep Neural Networks," *IEEE Antennas and Wireless Propagation Letters* 21, no. 8 (2023): 1562–1566.

5. Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel Estimation for Cell-Free mmWave Massive MIMO Through Deep Learning," *IEEE Transactions on Vehicular Technology* 68, no. 10 (2021): 10325–10329.

6. A. Abdallah, A. Celik, M. M. Mansour, and A. M. Eltawil, "Deep Learning-Based Frequency-Selective Channel Estimation for Hybrid mmWave MIMO Systems," *IEEE Transactions on Wireless Communications* 21, no. 6 (2021): 3804–3821.

7. K. J. Jang, S. Park, J. Kim, et al., "Path Loss Model Based on Machine Learning Using Multi-Dimensional Gaussian Process Regression," *IEEE Access* 10 (2023): 115061–115073.

8. P. K. Chundi, X. Wang, and M. Seok, "Channel Estimation Using Deep Learning on an FPGA for 5G Millimeter-Wave Communication Systems," *IEEE Transactions on Circuits and Systems I: Regular Papers* 69, no. 2 (2021): 908–918.

9. H. Cheng, S. Ma, and H. Lee, "CNN-Based mmWave Path Loss Modeling for Fixed Wireless Access in Suburban Scenarios," *IEEE Antennas and Wireless Propagation Letters* 19, no. 10 (2020): 1694–1698.

10. S. A. Aldossari, "Predicting Path Loss of an Indoor Environment Using Artificial Intelligence in the 28-GHz Band," *Electronics* 12, no. 3 (2023): 497.

11. T. S. Cousik, V. K. Shah, T. Erpek, Y. E. Sagduyu, and J. H. Reed, "Deep Learning for Fast and Reliable Initial Access in AI-Driven 6G mmWave Networks," *IEEE Transactions on Network Science and Engineering* 11 (2022): 5668–5680.

12. H. Shakhatreh, W. Malkawi, A. Sawalmeh, M. Almutiry, and A. Alenezi, "Modeling Ground-to-Air Path Loss for Millimeter Wave UAV Networks," arXiv preprint arXiv:2101.12024 (2021).

13. V. T. Quang, V. V. Yem, and H. T. P. Thao, "A Novel Method for Estimating Propagation Pathloss in Millimeter-Wave Communication Systems," *Bulletin of Electrical Engineering and Informatics* 12, no. 6 (2023): 3517–3529.

14. H. Li, X. Chen, K. Mao, et al., "Air-to-Ground Path Loss Prediction Using Ray Tracing and Measurement Data Jointly Driven DNN," *Computer Communications* 196 (2022): 268–276.

15. R.-T. Juang, "Deep Learning-Based Path Loss Model in Urban Environments Using Image-to-Image Translation," *IEEE Transactions on Antennas and Propagation* 70, no. 12 (2022): 12081–12091.

16. C. Sudhamani, M. Roslee, L. L. Chuan, A. Waseem, A. F. Osman, and M. H. Jusoh, "Enhanced Indoor Path Loss and RSRP of 5G mmWave Communication System With Multi-Objective Genetic Algorithm," *Wireless Personal Communications* 138, no. 1 (2024): 603–621.

17. M. T. Mezaal, N. B. M. Aripin, N. S. Othman, and A. H. Sallomi, "The Effect of Urban Environment on Large-Scale Path Loss Model's Main Parameters for mmWave 5G Mobile Network in Iraq," *Open Engineering* 14, no. 1 (2024) 20220601.

18. J. O. Afape, A. A. Willoughby, M. E. Sanyaolu, et al., "Improving Millimetre-Wave Path Loss Estimation Using Automated Hyperparameter-Tuned Stacking Ensemble Regression Machine Learning," *Results in Engineering* 22 (2024): 102289.

19. H. Zakeri, P. Khoddami, G. Moradi, et al., "Path Loss Model Estimation at Indoor Offices Environment by Using Deep Neural Network and CatBoost for Millimeter Wave 5G Wireless Application," *IEEE Access* 12 (2024): 159070–159085.

20. M. Brata and I. Zakia, "Path Loss Estimation at Sub-6 GHz and Millimeter Wave Frequencies Using Fine-Tuning," *IEEE Access* 12 (2024): 138142–138154.

21. S. Sung, W. Choi, H. Kim, and J.-I. Jung, "Deep Learning-Based Path Loss Prediction for Fifth-Generation New Radio Vehicle Communications," *IEEE Access* 11 (2023): 75295–75310.

22. J. Thrane, D. Zibar, and H. L. Christiansen, "Model-Aided Deep Learning Method for Path Loss Prediction in Mobile Communication Systems at 2.6 GHz," *IEEE Access* 8 (2020): 7925–7936.

23. S. Pawar and M. Venkatesan, "A Novel Pathloss Prediction and Optimization Approach Using Deep Learning in Millimeter Wave Communication Systems," *e-Prime Advances in Electrical Engineering, Electronics and Energy* 9 (2024): 100737.

24. K. Kayaalp, S. Metlek, A. Genc, H. Dogan, and İ. B. Basyigit, "Prediction of Path Loss in Coastal and Vegetative Environments With Deep Learning at 5G Sub-6 GHz," *Wireless Networks* 29, no. 6 (2023): 2471–2480.

25. H. Cheng, S. Ma, H. Lee, and M. Cho, "Millimeter Wave Path Loss Modeling for 5G Communications Using Deep Learning With Dilated Convolution and Attention," *IEEE Access* 9 (2021): 62867–62879.

26. S. E. Hadji and M. Nedil, "Machine learning-based underground mine path loss prediction using mm-Wave massive MIMO measurements," *IEEE Transactions on Antennas and Propagation* 73 (2025): 9378–9390.

27. S. Hussain, "A Multi-Scale Feature Extraction and Fusion UNet for Pathloss Prediction in UAV-Assisted mmWave Radio Networks," arXiv preprint arXiv:2509.09606 (2025).

28. S. A. Robinson, A. A. Imianvan, E. C. Igodan, E. A. Dan, K. U. Joseph, and L. A. Dickson, "Machine Learning Approach for Path Loss Prediction in Urban Drive 5G Network Environments," *International Journal of Microwave & Optical Technology* 20, no. 4 (2025): 380–388.

29. J.-S. Hong, I. Hermann, F. G. Zöllner, et al., "Acceleration of Magnetic Resonance Fingerprinting Reconstruction Using Denoising and Self-Attention Pyramidal Convolutional Neural Network," *Sensors* 22, no. 3 (2022): 1260.

30. S. Xiao, D. Zhu, C. Tang, and Z. Huang, "Combining Graph Contrastive Embedding and Multi-Head Cross-Attention Transfer for Cross-Domain Recommendation," *Data Science and Engineering* 8, no. 3 (2023): 247–262.

31. T. Hamadneh, K. Kaabneh, O. AlSayed, G. Bektemyssova, Z. Montazeri, and M. Dehghani, "Sculptor Optimization Algorithm: A new Human-Inspired Metaheuristic Algorithm for Solving Optimization Problems," *International Journal of Intelligent Engineering & Systems* 17, no. 4 (2024): 564–575.

32. S. Reza, M. C. Ferreira, J. J. M. Machado, and J. M. R. S. Tavares, "A Customized Residual Neural Network and Bi-Directional Gated Recurrent Unit-Based Automatic Speech Recognition Model," *Expert Systems with Applications* 215 (2023): 119293.

33. M. Dehghani, Š. Hubálovský, and P. Trojovský, "Northern Goshawk Optimization: A New Swarm-Based Algorithm for Solving Optimization Problems," *IEEE Access* 9 (2021): 162059–162080.

34. S. Alomari, K. Kaabneh, I. AbuFalahah, et al., "Carpet Weaver Optimization: A Novel Simple and Effective Human-Inspired Metaheuristic Algorithm," *International Journal of Intelligent Engineering & Systems* 17, no. 4 (2024): 230–242.

35. H. Zhang, Y. Ma, K. Yuan, M. Khayatnezhad, and N. Ghadimi, "Efficient Design of Energy Microgrid Management System: A Promoted Remora Optimization Algorithm-Based Approach," *Heliyon* 10, no. 1 (2024): e23394.

36. H. Al-Khazraji, A. R. Nasser, A. M. Hasan, A. K. al Mhdawi, H. Al-Raweshidy, and A. J. Humaidi, "Aircraft Engines Remaining Useful Life Prediction Based on a Hybrid Model of Autoencoder and Deep Belief Network," *IEEE Access* 10 (2022): 82156–82163.

37. A. A. Mohammed and P. Sumari, "Hybrid K-Means and Principal Component Analysis (PCA) for Diabetes Prediction," *International Journal of Computing and Digital Systems* 15, no. 1 (2024): 1719–1728.

38. D. Thakur, S. Biswas, E. S. L. Ho, and S. Chattopadhyay, "ConVAE-LSTM: Convolutional Autoencoder Long Short-Term Memory Network for Smartphone-Based Human Activity Recognition," *IEEE Access* 10 (2022): 4137–4156.