一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
A・P2026-13(2026-04)

Language Model-Empowered Transformer Framework for Antenna Inverse Design

Chang GE† and Qiang CHEN†

†School of Engineering, Tohoku University
6-6-05, Aramaki Aza Aoba Aoba-ku, Sendai, Miyagi 980-8579, Japan

E-mail: ge.chang.e5@tohoku.ac.jp, qiang.chen.a5@tohoku.ac.jp

Abstract Antenna design is fundamentally an inverse problem that maps desired electromagnetic performance to geometric structures and design parameters. Conventional approaches typically rely on repeated full-wave computational electromagnetics (CEM) simulations for iterative search and optimization, resulting in high computational cost and long design cycles. Motivated by recent advances in large language models (LLMs), this paper introduces an LLM-inspired framework for antenna inverse design, where the language "translation" paradigm is extended from time-ordered sequences to spatially structured sequences tailored to antenna representations. This perspective provides a promising and extensible pathway toward efficient end-to-end antenna inverse design, with strong potential for future research and practical engineering applications.

Key words Antenna Inverse Design, Transformer, Large Language Model, Yagi-Uda antennas.

1. Introduction

Antenna design is fundamentally an inverse problem, namely, mapping desired electromagnetic performance back to concrete geometric structures and design parameters. Most existing approaches rely on repeatedly invoking full-wave computational electromagnetics (CEM) solvers for iterative search and optimization, which is computationally expensive and significantly prolongs the design cycle. If an end-to-end antenna inverse design workflow can be realized through a single forward pass of a state-of-the-art AI model, repeated full-wave simulations during the design stage would no longer be necessary. Such a paradigm could markedly improve design efficiency, thereby accelerating both large-scale exploration and practical deployment.

Considerable progress has been made toward this goal in recent years. However, most existing studies are still built upon relatively conventional architectures, such as multilayer perceptrons (MLPs) [1], convolutional neural networks (CNNs) [2], and recurrent neural networks (RNNs) [3]. In contrast, the rapid rise of large language models (LLMs) has recently reshaped the landscape of artificial intelligence and redefined its development trajectory. With their remarkable capabilities in representation learning, sequence modeling, and complex reasoning, LLMs are driving many disciplines to seek new points of connection with the emerging LLM ecosystem, as evidenced by recent efforts in various fields [4]-[7].

Through our recent studies, we find that the sequence modeling paradigm represented by large language models (LLMs) is naturally well aligned with antenna problems, making antenna design an excellent testbed for LLM-inspired engineering applications. Based on this insight, this paper introduces a new direction for antenna inverse design by incorporating LLM-inspired ideas into antenna design. Specifically, drawing on the "translation" paradigm in natural language processing, we employ the Transformer architecture [8], which underlies representative LLMs such as ChatGPT and Gemini, to bridge LLM-style sequence modeling with antenna inverse design. To the best of the authors' knowledge, this work constitutes the first attempt to connect antenna design with the LLM paradigm. This perspective opens an important opportunity to rethink existing AI frameworks for antenna inverse design and to explore whether LLM-inspired modeling can provide a more powerful and extensible solution, with strong potential for future research and practical engineering applications.

2. Dataset Construction

The dataset used in this work is generated by our in-house full-wave simulator, V-CALAR, which implements a method-of-moments (MoM) formulation based on Richmond's thin-wire modeling approach. In this study, Yagi–Uda arrays are chosen as the benchmark antenna structure. All simulations are performed in a wavelength-normalized framework, where all geometric dimensions are expressed as dimensionless ratios to the wavelength. Therefore, no explicit design frequency f_0 is specified at the solver level, and the resulting dimensionless antenna designs can be scaled to any target frequency band during practical implementation.

In total, the dataset contains 83,877 samples for 2-element Yagi–Uda arrays and 3,686,575 samples for 3-element arrays. All geometric parameters are randomly sampled within the ranges summarized in Table I. Assuming a sampling resolution of 0.01, the generated dataset corresponds to only about 0.3% of the full design space obtained by discretizing all parameters in Table I with a step size of 0.01. This indicates that, despite the seemingly large number of samples, the dataset still covers only a small portion of the highly complex design space. Generating all samples with a single CPU core required approximately one week of wall-clock time and produced about 400 GB of data.

For data splitting, 1,206 samples from the 2-element dataset and 39,066 samples from the 3-element dataset are randomly reserved as the test set. The remaining samples are further divided into training and validation sets with a ratio of 80:20. As a result, the final training set contains 66,137 samples for the 2-element case and 2,918,007 samples for the 3-element case.

TABLE I

RANGES OF GEOMETRICAL PARAMETERS (NORMALIZED BY WAVELENGTH)

Antenna configuration	Parameters [λ]	Value range
2-element	l_1	0.1-1.0 λ
	d_1	0.1-0.5 λ
	l_2	0.1-1.0 λ
3-element	l_1	0.1-1.0 λ
	d_1	0.1-0.5 λ
	l_2	0.1-1.0 λ
	d_2	0.1-0.5 λ
	l_3	0.1-1.0 λ

3. Model Design

We adopt a classical encoder–decoder Transformer architecture in this work. For the encoder, the input directivity pattern is encoded in a manner similar to patch tokens with positional information [9], so as to capture its global spatial dependencies.

For the decoder, the standard Transformer structure is modified to accommodate the multi-configuration characteristics of Yagi–Uda antennas. Specifically, we design a multi-head output scheme to more appropriately adapt to different antenna configurations within a unified framework. In addition, the standard autoregressive decoding strategy is replaced with a teacher-forcing-based parallel output mechanism. This modification makes the model more suitable for antenna design, because the target parameters are fixed-dimensional continuous variables and do not naturally require token-by-token autoregressive generation as in language modeling.

Furthermore, since antenna parameter prediction is essentially a regression task rather than a classification task, the softmax-based classification head in the standard Transformer is replaced with linear output layers to directly predict continuous geometric parameters. The model is trained by minimizing the mean squared error

(MSE) between the predicted and ground-truth antenna parameters.

4. Performance Evaluation

To evaluate the accuracy performance of prediction, we examine the forward and backward directivity of the realized 2-element and 3-element Yagi–Uda arrays. The corresponding error statistics are summarized in Table II, showing that the prediction errors remain low for both configurations. In addition, Figs. 1 present representative E-plane and H-plane results, where the target patterns are compared with the realized patterns. A close agreement can be observed between the two, with only minor discrepancies. Taken together, these results demonstrate that the Transformer can effectively learn the mapping from target antenna performance to geometric design parameters.

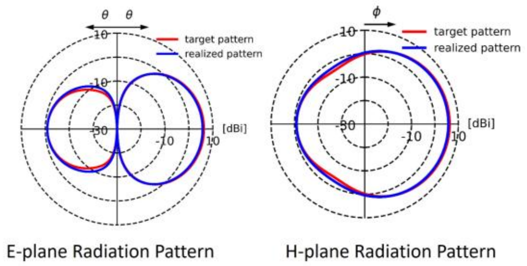
TABLE II

ERRORS OF FORWARD AND BACKWARD DIRECTIVITY

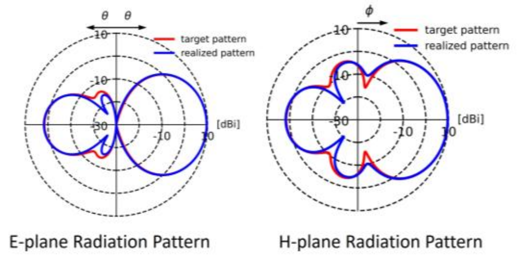
Antenna configuration	Direction	Directivity MAE (dB)	Directivity RMSE (dB)	Directivity STD (dB)
2-element	Forward	0.014	0.071	0.070
	Backward	0.016	0.147	0.147
3-element	Forward	0.086	0.300	0.300
	Backward	0.083	0.334	0.334

5. Conclusion

In this work, we explored the use of the Transformer architecture, which serves as the foundation of modern large language models (LLMs), for antenna inverse design. By drawing an analogy to language translation, the proposed framework learns a high-dimensional mapping from desired directivity patterns to antenna geometric parameters. Experimental results on Yagi–Uda antennas demonstrate promising prediction accuracy and validate the effectiveness of this approach. Overall, this study opens a new direction for antenna inverse design and highlights the potential of integrating advanced AI models with physics-driven electromagnetic modeling.



Target parameters: $l_1 = 0.45\lambda$, $d_1 = 0.15\lambda$, $l_2 = 0.42\lambda$
Predicted parameters: $l_1' = 0.42\lambda$, $d_1' = 0.16\lambda$, $l_2' = 0.41\lambda$



Target parameters: $l_1 = 0.48\lambda$, $d_1 = 0.26\lambda$, $l_2 = 0.46\lambda$, $d_2 = 0.20\lambda$, $l_3 = 0.44\lambda$
Predicted parameters: $l_1' = 0.47\lambda$, $d_1' = 0.27\lambda$, $l_2' = 0.45\lambda$, $d_2' = 0.19\lambda$, $l_3' = 0.44\lambda$

Fig.1. Comparison between the target and realized E-plane and H-plane directivity cuts for a representative 2/3-element Yagi–Uda array.

6. Acknowledgment

This work was partially supported by the supporting program from Center for Diversity, Equity and inclusion (TUMUG) at Tohoku University.

7. References

[1] J. Zhang, K. Jin, E. Zhu and W. -Y. Yin, "Inverse Design of Freely-Partitioned Metasurface Antennas via Simulation-Driven Active Learning and Deep Neural Network," in *IEEE Antennas and Wireless Propagation Letters*, doi: 10.1109/LAWP.2025.3621829.

[2] A. Gupta, E. A. Karahan, C. Bhat, K. Sengupta and U. K. Khankhoje, "Tandem Neural Network Based Design of Multiband Antennas," in *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 8, pp. 6308-6317, Aug. 2023, doi: 10.1109/TAP.2023.3276524.

[3] T. Yuan, J. Xie, Y. Li, X. Yang, X. Yang and E. Wen, "Flex-Form Generative Neural Network for Multifrequency Modes Antennas," in *IEEE Transactions on Antennas and Propagation*, vol. 73, no. 10, pp. 8313-8318, Oct. 2025, doi: 10.1109/TAP.2025.3584591.

[4] E. Purificato *et al.*, *The Role of Artificial Intelligence in Scientific Research: A Science for Policy, European Perspective*. Luxembourg, Luxembourg: Publications Office of the European Union, 2025. [Online]. Available: https://data.europa.eu/doi/10.2760/7217497

[5] X. Jiang, D. Xue, Y. Bai, W. Y. Wang, J. Liu, M. Yang, and Y. Su, "AI4Materials: Transforming the landscape of materials science and engineering," *Rev. Mater. Res.*, vol. 1, art. no. 100010, 2025, doi: 10.1016/j.revmat.2025.100010.

[6] T. L. Khoo, T. S. Lee, S.-T. Bee, C. Ma, and Y.-Y. Zhang, "A comparative review of large language models in engineering with emphasis on chemical engineering applications," *Processes*, vol. 13, no. 9, art. no. 2680, Aug. 2025, doi: 10.3390/pr13092680.

[7] J. Guo *et al.*, "Large language model-empowered next-generation computer-aided engineering," *arXiv:2509.11447*, 2025.

[8] A. Vaswani et al., "Attention Is All You Need," https://doi.org/10.48550/arXiv.1706.03762

[9] A. Dosovitskiy et al., "An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale," https://doi.org/10.48550/arXiv.2010.11929