

Optimal Neural Network Hyperparameter Consideration through Digital Predistortion in RF Power Amplifiers

Original

Optimal Neural Network Hyperparameter Consideration through Digital Predistortion in RF Power Amplifiers / Kouhalvandi, L., Matekovits, L., Aygun, S., Ozoguz, S.. - ELETTRONICO. - (2026), pp. 1-4. (2026 34th Signal Processing and Communications Applications Conference (SIU) Istanbul (Tur) 07-10 July 2026) [10.1109/siu71813.2026.11636623].

Availability:

This version is available at: 11583/3015151 since: 2026-09-03T13:44:25Z

Publisher:

IEEE

Published

DOI:10.1109/siu71813.2026.11636623

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

IEEE postprint/Author's Accepted Manuscript

©2026 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

RF Güç Yükselteçlerinde Dijital Ön Bozulma Yöntemiyle Optimal Sinir Ağı Hiperparametrelerinin Değerlendirilmesi

Optimal Neural Network Hyperparameter Consideration through Digital Predistortion in RF Power Amplifiers

Authors Blinded

Özetçe —Bu çalışma, güç amplifikatörlerinin (PA) davranışını modellemek için tasarlanmış derin bir sinir ağının (DNN) en uygun hiperparametrelerini belirlemek için etkili bir optimizasyon tekniği sunmaktadır. Bu nedenle, PA'ların hem modelleme doğruluğunu hem de dijital ön bozulma (DPD) performansını iyileştirmek için yenilikçi ve verimli bir verim analizi tabanlı yaklaşım önerilmiştir. Yöntem, PA sistemlerinde bulunan doğrusal olmayan ve bellek etkilerini yakalamak için uzun kısa süreli bellek (LSTM) DNN mimarisinden yararlanarak genel performansı ve doğrusallaştırma yeteneğini artırmaktadır. En uygun eğitimi elde etmek için, öğrenme oranı, ağ derinliği ve nöron konfigürasyonu gibi en uygun hiperparametreleri belirlemek üzere sistematik olarak birden fazla optimizasyon stratejisi uygulanmaktadır. Önerilen yöntemin etkinliğini doğrulamak için, 1,8 GHz ile 2,2 GHz frekans aralığında çalışan bir PA ele alınmıştır ve kapsamlı simülasyonlar ve değerlendirmeler, optimize edilmiş DNN'nin minimum modelleme hatası elde ederken DPD performansını önemli ölçüde iyileştirdiğini göstermektedir. Sonuçlar, önerilen çerçevenin PA modellemesi ve doğrusallaştırılması için güvenilir ve verimli bir çözüm sunduğunu ve doğruluk ve tutarlılık açısından geleneksel tekniklerden daha iyi performans gösterdiğini doğrulamaktadır.

Anahtar Kelimeler—Dijital ön bozulma (DPD), derin sinir ağı (DNN), hiperparametre, Güç yükselteci (PA), optimizasyon.

Abstract—This study introduces an effective optimization technique for identifying the optimal hyperparameters of a deep neural network (DNN) designed to model the behavior of power amplifiers (PAs). Hence, an innovative and efficient yield-analysis-based approach is proposed to improve both the modeling accuracy and the digital predistortion (DPD) performance of PAs. The method leverages a long short-term memory (LSTM) DNN architecture to capture the nonlinear and memory effects inherent in PA systems, thereby enhancing overall performance and linearization capability. To achieve optimal training, multiple optimization strategies are systematically applied to determine the most suitable hyperparameters, such as learning rate, network depth, and neuron configuration. To validate the effectiveness of the proposed method, a PA operating in the 1.8 GHz up to 2.2 GHz frequency range is considered for which extensive simulations and evaluations demonstrate that the optimized DNN achieves minimal modeling error while significantly improving DPD performance. The results confirm that the proposed framework offers a reliable and efficient solution for PA modeling and linearization, outperforming conventional techniques in terms of

accuracy and consistency.

Keywords—Digital predistortion (DPD), deep neural network (DNN), hyperparameter, power amplifier (PA), optimization.

I. INTRODUCTION

Power amplifiers (PAs) are essential components in modern wireless communication systems. Although they achieve high efficiency near saturation, PAs exhibit strong nonlinear behavior, which can degrade signal quality, including increases in adjacent channel power ratio (ACPR). To mitigate these effects, digital predistortion (DPD) techniques are commonly used to linearize PA output. At the same time, wireless communication systems are evolving toward greater intelligence, enabling radio frequency (RF) transceiver configurations to be dynamically adjusted based on factors such as user demand, distance, and transmission direction [1]–[3]. Recently, the application of neural networks (NNs) have proved their effectiveness in the DPD domain [4], [4], [5].

In [6], a cascade deep neural network (DNN) framework to compensate for a wide range of RF hardware impairments in wireless transmitters, extending beyond conventional approaches that mainly address PA nonlinearities is presented. A Cascade NN-based DPD model that targets multiple distortion sources across the RF chain, including digital-to-analog converter nonlinearity, in-phase/quadrature (IQ) imbalance, and PA effects is introduced. In another study, [7], a convolutional neural network (CNN)-based multi-beam DPD technique for phased-array transmitters, designed to address nonlinear distortions, hardware impairments, and cross-beam coupling effects across a wide range of beam steering angles is presented. Unlike conventional methods that require interpolation, the presented approach incorporates steering angle information together with I/Q signals from both the main and neighboring subarrays, enabling accurate compensation for both in-band and out-of-band distortions over varying beam configurations. In [8], Kronecker product-based module to enhance NN-based DPD for wideband PAs, addressing the high complexity typically required by conventional models is presented. The method introduces a structured way of generating nonlinear features by combining linear layers with Kronecker product

operations, inspired by memory polynomial representations. This allows the network to capture nonlinear behavior more efficiently while reducing the number of required parameters. In another study, [9], a DPD method for short-wave high-frequency PAs using an enhanced echo state network to jointly address in-band intermodulation distortion (IMD) and out-of-band harmonic distortion is presented. Traditional DPD techniques primarily target IMD, while harmonic components require external analog filters due to their frequencies exceeding the DPD sampling bandwidth. Also, in [10], a cross-architecture knowledge distillation method for DPD of terahertz and millimeter-wave transceivers, aimed at improving the linearization of PAs under resource-constrained deployment is presented.

This paper devotes to investigate various optimization methods to identify the most suitable approach for determining the optimal hyperparameters of a DNN used for modeling PA behavior. It focuses on optimizing key hyperparameters, including the neuron number (NN), layer number (LN), learning rate (LR), dropout rate (DR), and batch size (BS). In addition, a regression-based DNN incorporating long short-term memory (LSTM) layers is employed to enhance DPD performance. The proposed framework provides an effective solution for PA modeling and facilitates the selection of optimal hyperparameters, thereby improving the accuracy and efficiency of behavioral modeling through the utilization of DPD outcome. For this case, various optimization methods such as genetic algorithm (GA) [11], particle swarm optimization (PSO) [12], whale optimization algorithm (WOA) [13], thompson sampling efficient multiobjective optimization (TSEMO) [14], and multi-objective ant lion optimizer (MOALO) [15] are executed. These methods are employed for modeling a PA with DPD specification operating with 400 MHz bandwidth for proving the effectiveness of method. It is decided that MOALO is suitable enough in training the DNN in which the hyperparameters are determined through this algorithm leading to train an accurate DNN.

This paper is organized as follows: Section II presents the proposed methodology leading to achieve optimal hyperparameters of trained DNN. Section III describes the simulation results and Section IV concludes this work.

II. PROPOSED METHODOLOGY

The DNNs, particularly multi-layer architectures, have demonstrated strong performance and high accuracy in various RF applications. Motivated by this capability, a regression-based DNN incorporating LSTM layers is employed to model a practical PA while accounting for its DPD characteristics over a given frequency band. This section presents the proposed methodology for determining the optimal hyperparameters required for accurate PA modeling and effective DPD performance.

To ensure effective NN performance, it is essential to properly select key hyperparameters, including the LN, LR, DR, BS, and activation functions, as these directly influence the behavior of the trained model. In this study, several optimization methods are explored and applied to identify the optimal hyperparameter configuration. A comparative evaluation is then conducted to determine the most suitable method for

training LSTM-based DNNs with respect to accuracy. Figure 1 illustrates the overall architecture of the proposed DNN used to model the DPD performance of a practical PA based on I and Q components. Achieving effective training requires the selection of appropriate model hyperparameters. In this work, several optimization techniques are employed, including the GA, PSO, WOA, TSEMO, and MOALO algorithms and train the DNN to consider the accuracy of the DNN.

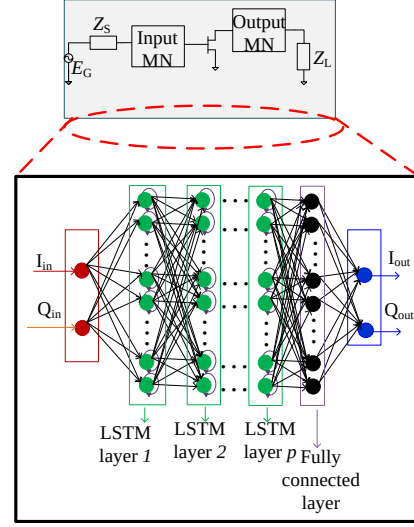


Figure 1: PA behavioral modeling through the LSTM-based DNN.

The first and most critical step in training the DNN is the preparation of an adequate dataset, including training (X_{Train}), validation (X_{Val}), and testing (X_{Test}) inputs, along with their corresponding target outputs (Y_{Train} , Y_{Val} , and Y_{Test}). In this work, the dataset is divided using a ratio of 70%, 15%, and 15%, respectively. The data are generated through a co-simulation framework that integrates Keysight ADS with a numerical analysis tool such as MATLAB. Design parameters of the PA are varied within a range of $\pm 10\%$ to $\pm 30\%$ with a step size of $\pm 5\%$. The resulting dataset consists of in-phase and quadrature components, denoted as I_i and Q_i for $i = 0, 1, \dots, k$, corresponding to different PA parameter configurations.

Following dataset generation, the input and output features are defined, and the regression DNN is constructed using LSTM-based hidden layers along with fully connected layers. Subsequently, the hyperparameters are specified and optimized using various optimization techniques. The performance of the model is evaluated by comparing the actual testing outputs (Y_{Test}) with the predicted outputs (Y_{Pred}). The *'predictAndUpdateState'* function is then utilized for accurate PA modeling. Model convergence is assessed using the normalized root mean square error (RMSE) on the testing data, which is reported as the testing accuracy and is expected to be less than 1. In this framework, the rectified linear unit (ReLU) is adopted as the activation function, while the mean squared error (MSE) is used as the loss function for training the regression DNN.

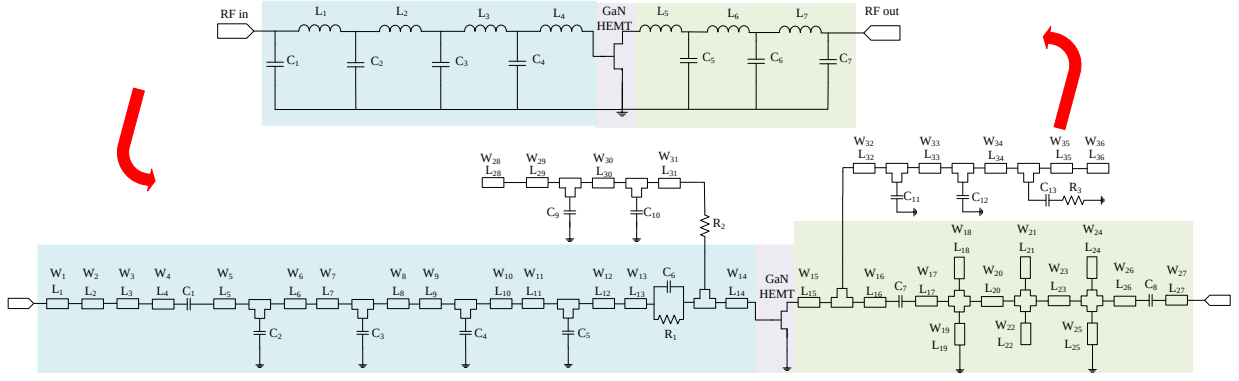


Figure 2: PA structure presented in [16].

Table I: PA modeling with DPD specification using various optimization methods

Method	NN	LN	LR	DR	BS
GA	56	1	0.02	0.3	0.8
PSO	150	2	0.01	0.5	1
WOA	130	2	0.01	0.5	1
TSEMO	200	2	0.006	0.5	0.8
MOALO	300	5	0.005	0.5	1

III. SIMULATION RESULTS

To validate the proposed approach, an appropriate dataset is collected based on the PA structure illustrated in Figure 2. The simulations are carried out using a computational setup consisting of an Intel Core i7-4790 CPU operating at 3.60 GHz with 32 GB of RAM. The PA design employed in this study follows the configuration presented in [16], where detailed design parameters are provided. This section presents and discusses the simulation results obtained to evaluate the effectiveness of the proposed method.

As an initial step, 1000 sequences consisting of I and Q data are extracted from the established setup. The hidden layer architecture, as illustrated in Figure 1, is then defined. Subsequently, the training configuration is specified by selecting the Adam solver and setting the gradient threshold to 1. The network parameters, including weights and biases, are updated using the Adam optimization algorithm along with standard gradient descent. Next, the key hyperparameters—namely the NN, LN, LR, DR, and BS are defined as optimization variables. These parameters are optimized using the selected methods described in Section II, with the objective of achieving an optimal regression DNN tailored for DPD performance. Table I presents the various hyperparameters of trained DNN through optimization methods and Figure 3 presents the overall accuracy by employing the nominated algorithms. Also Figure 4 shows the RMSE performance that in 3th layer the normalized RMSE is below 0.1. Here with the utilization of MOALO method, the DNN specifications are as follows: NN=300, LN=5, LR=0.005, DR=0.5, BS=1. Additionally, the DPD performance for the PA is presented in Figure 5.

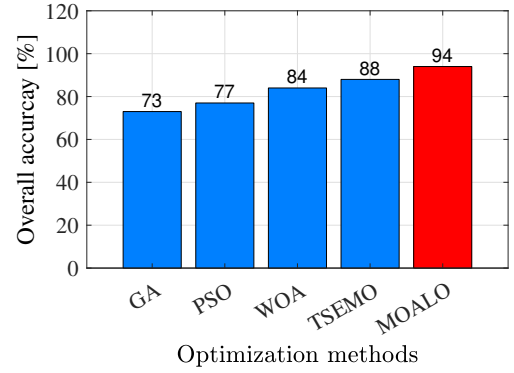


Figure 3: Comparison between various optimization methods for training the DNN.

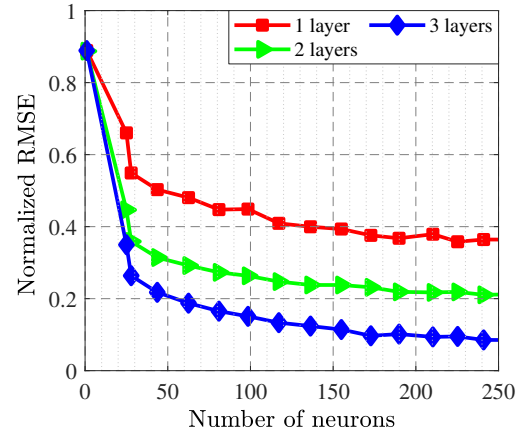


Figure 4: RMSE performance of trained LSTM-based DNN.

IV. CONCLUSION

This paper presents a methodology for modeling PAs using the DNN by optimizing the hyperparameters of an LSTM-based architecture. The proposed approach assists designers in constructing reliable DNN models while improving DPD performance, thereby significantly reducing design effort. Various optimization techniques are explored to determine key hyperparameters based on I and Q signal components. The DNN is trained and evaluated using PA data, and simulation

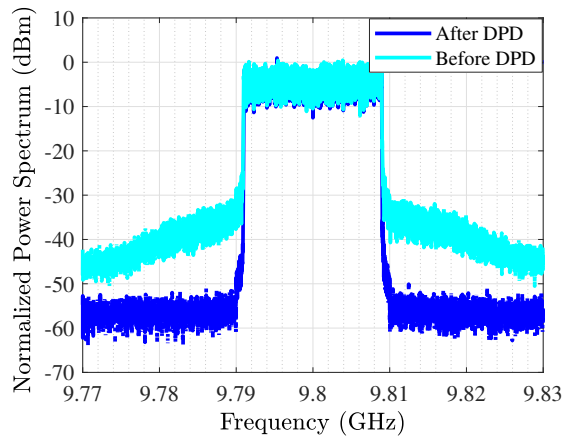


Figure 5: DPD performance.

results indicate that the MOALO method outperforms other techniques in identifying optimal hyperparameters for the LSTM-based DNN. This approach provides a practical and effective solution for PA designers, enabling more accurate behavioral modeling through optimized hyperparameter selection.

KAYNAKLAR

- [1] J. Huang, C. Yu, Z. Chen, Z. Zheng, and Y. Liu, "Dynamic convolutional neural network for digital predistortion in multistate pa modeling," *IEEE Microwave and Wireless Technology Letters*, pp. 1–4, 2026.
- [2] J. Chen, G. Xu, W. Zhao, Y. Lv, T. Liu, and F. M. Ghannouchi, "Exploring dynamic sparse memory effect under varying transmission states for digital predistortion of rf power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 74, no. 1, pp. 581–595, 2026.
- [3] Y. Zhang, Q. Chen, H. Peng, W. Chen, Z. Feng, and F. M. Ghannouchi, "Stream-level digital predistortion for extremely large-scale mimo transmitters," *IEEE Transactions on Microwave Theory and Techniques*, pp. 1–17, 2026.
- [4] Y. Wu, Q. Wang, T. Ohtsuki, D. Niyato, F. Adachi, and G. Gui, "Generalized wideband power amplifier modeling for vehicular wireless transceivers using kolmogorov-arnold convolutional liquid neural networks," *IEEE Transactions on Vehicular Technology*, pp. 1–6, 2026.
- [5] M. T. Khan, Y. Ding, and G. Goussetis, "Next-gen digital predistortion from hardware acceleration of neural networks: Trends, challenges, and future," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–21, 2026.
- [6] A. Mohammad, A. Fischer-Bühner, L. Anttila, M. Valkama, and B. Tan, "A cascade deep neural network approach for compensation of rf hardware impairments in wireless transmitters," *IEEE Open Journal of the Communications Society*, vol. 7, pp. 1249–1262, 2026.
- [7] M. H. Khazani, A. Motaqi, A. Dalbah, M. Helaloui, W. Chen, and F. M. Ghannouchi, "A cnn-based digital predistortion method for multi-beam phased-array transmitters," *IEEE Wireless Communications Letters*, vol. 15, pp. 56–60, 2026.
- [8] J. Shao, X. Hong, and W. Wang, "A kronecker product-based module for digital predistortion neural network for the wide bandwidth power amplifier," *IEEE Microwave and Wireless Technology Letters*, vol. 36, no. 1, pp. 143–146, 2026.
- [9] J. Ren, C. Wang, M. Wu, Y. Liu, and Y. Li, "An integrated scheme of enhanced echo state network of harmonic cancellation digital predistortion for short-wave power amplifiers," *IEEE Microwave and Wireless Technology Letters*, vol. 36, no. 1, pp. 139–142, 2026.
- [10] G. Zhao, K. Ying, L. Zhao, and L. Gui, "Cross-architecture knowledge distillation for digital predistortion of terahertz/mmwave transceiver," *IEEE Internet of Things Journal*, vol. 13, no. 3, pp. 5006–5017, 2026.
- [11] M. Yan, Z. Long, B. Lin, and Y. Liu, "A genetic algorithm–sparse bayesian joint optimization modeling approach for efficient predistortion," *IEEE Microwave and Wireless Technology Letters*, vol. 35, no. 11, pp. 1668–1671, 2025.
- [12] R. Ling, Z. Zhang, and X. Xuan, "Dual-frequency power amplifiers' design based on improved multiobjective particle swarm optimization algorithm," *IEEE Microwave and Wireless Technology Letters*, vol. 35, no. 11, pp. 1768–1771, 2025.
- [13] M. I. Habelalmateen, R. Archana Reddy, C. Reddy, N. Nagarjuna, and P. K. S., "Papr reduction in mimo-ofdm using nonlinear inertia weight factor of whale optimization algorithm," in *2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)*, 2024, pp. 1–5.
- [14] =plus 4minusE.Bradford,A.M.Schweidtmann,andA.Lapkin, "Efficientmultiobjectiveoptimizationemployinggussianprocesses, spectralsamplingandageneticalgorithm," *Journal of Global Optimization*, vol. 71, no. 2, pp. 407–438, Jun 2018. [Online]. Available: <https://doi.org/10.1007/s10898-018-0609-2>0pt
- [15] =plus4minusS.Mirjalili,P.Jangir,andS.Saremi, "Multi-objectiveantlion optimizer:amulti-objectiveoptimizationalgorithmforsolvingengineering problems," *Applied Intelligence*, vol. 46, no. 1, pp. 79–95, Jan 2017. [Online]. Available: <https://doi.org/10.1007/s10489-016-0825-8>=0pt
- [16] =plus4minusA.BLINDED, "Authorsblinded," *Sensors*, vol. 25, no. 0, 2025. [Online]. Available: 0=0pt