Improved Temperature- and Power-Dependent Convolutional NN-Based PA

José D. Domingues, *Graduate Student Member, IEEE*, André Prata, and Jiri Stulemeijer, *Senior Member, IEEE*

***Abstract*— This letter presents an innovative power amplifier (PA) behavioral model (BM) method valid for a range of different ambient temperatures and input power levels. This work presents a novel input image layer for a real-valued time-delay convolutional neural network (RVTDCNN). This image layer uses preprocessed ambient temperature and dissipated power. The preprocessed temperature and power as well as the present samples are placed in a central position inside the image layer. This maximizes the number of convolution operations that they are included in thereby magnifying the importance of these inputs in the feature maps. The newly proposed method delivers, in comparable conditions, a normalized mean square error (NMSE) improvement of over 3 dB compared to a previously published method.**

***Index Terms*— Artificial neural network (NN), linear and nonlinear device modeling, power amplifier (PA).**

I. Introduction

FOR the sake of increasing efficiency, power amplifiers (PAs) are usually operated in a nonlinear regime. Operating in this regime degrades signal quality metrics, such as error vector magnitude (EVM) and adjacent channel leakage ratio (ACLR). The prevailing technique to counteract this degradation is digital predistortion (DPD), by precorrecting for the PA nonlinear behavior.

DPD investigation and fine-tuning can benefit from a highly accurate behavioral model (BM). The PA behavior depends on several parameters, such as bias, frequency, temperature, and power. This work focuses on these last two.

While there are several traditional linearization and BM techniques based on memory polynomials [1], [2], lookup tables [3], [4], and others, these usually work for only one PA state. Neural networks (NNs) are powerful candidates to implement BMs. Not only do they excel at approximating complex nonlinear functions [7], [8], but they also yield the opportunity to construct a parameterized BM [5], [6]. By using a convolutional layer to capture in an effective manner, the important features of the input data, the size of the network can be kept relatively small. This is possible due to the weight sharing and data dimensionality reduction characteristic of the convolutional layer as well as strong modeling capabilities of the convolutional NN (CNN) [5].

Recently, a real-valued time-delayed CNN (RVTDCNN)-based architecture for PA BM was presented for wideband PA modeling [6]. This model achieved high-performance PA modeling and linearization for signals with bandwidths higher than 100 MHz while achieving low complexity and fast convergence. However, it was made for a single PA state.

In [5], power–temperature inclusive (PTI)-DPD was developed. This RVTDCNN model-based DPD is capable of linearizing over PA states. The model is parameterized over a range of ambient temperatures and input power back-offs (IPBOs) [5].

This letter improves over the state of the art by preprocessing the input data fed into the RVTDCNN-based architecture developed in [5] and changing the image layer format, resulting in improved model performance while maintaining the same network complexity. Model performance has been evaluated in terms of normalized mean square error (NMSE) and ACLR. This letter is structured as follows. Section II describes the architecture of the network used to implement the BM as well as the innovations made upon it. Section III presents the results obtained for the newly proposed model. Finally, in Section IV the conclusion of this work is presented.

II. NN Architecture and Improvements

A. Neural Network Architecture

The implemented network architecture for the BM is presented in Fig. 1. From the left to right, we have a flow that starts in the input layer with the input data, *IQ* input data ($x(n)$), ambient temperature T, and signal average power P. The input data are fed into the network. The PA model output is the *IQ* sample on the right. The input layer processes the data input points in a way to fill a matrix with size 5×3. Each matrix will include the present time waveform sample, as well as a lead and a lag sample. The matrix will be filled in the first and second rows with I and Q samples, respectively.

The third row is filled with the PA operating conditions, i.e., the preprocessed temperature and power data referring to the state of the PA to be modeled.

The last two rows are used to carry envelope-dependent terms of the signal $|x(n)| = \text{sqrt}(I(n)\text{^}2 + Q(n)\text{^}2)$ and its memory parts having two nonlinearity orders, as shown in Fig. 2. As stated in [5], these nonlinearity orders will combine in the convolutional layer to form higher order nonlinearities.

The utilization of a convolutional layer in the network design enhances the efficiency of the network by effectively retaining the key features of the input data while reducing its

Manuscript received 28 February 2023; revised 12 April 2023; accepted 14 April 2023. This work was supported in part by the Fundação para a Ciência e Tecnologia (FCT) under Ph.D. Grant SFRH/BD/07051/2021 and in part by FCT/Ministério da Ciência, Tecnologia e EnsinoSuperior (MCTES) through national funds and when applicable co-funded European Union (EU) funds under Project UIDB/50008/2020-UIDP/50008/2020. *(Corresponding author: José D. Domingues.)*

José D. Domingues is with the Instituto de Telecomunicações, 3810-193 Aveiro, Portugal (e-mail: j.domingues@ua.pt).

André Prata and Jiri Stulemeijer are with Qualcomm Technologies Netherlands B.V., 6546 AS Nijmegen, The Netherlands (e-mail: aprata@qti.qualcomm.com; jiris@qti.qualcomm.com).

This article was presented at the IEEE MTT-S International Microwave Symposium (IMS 2023), San Diego, CA, USA, June 11–16, 2023.

Color versions of one or more figures in this letter are available at https://doi.org/10.1109/LMWT.2023.3268796.

Digital Object Identifier 10.1109/LMWT.2023.3268796

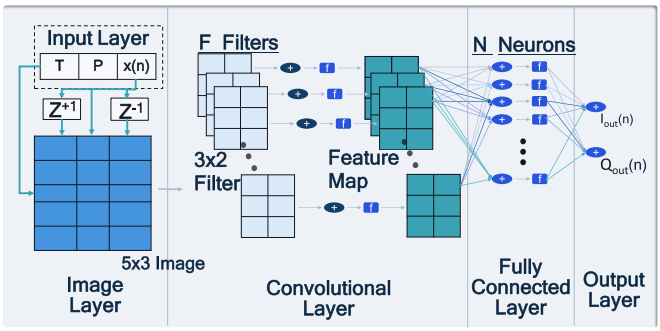


Fig. 1. Network architecture, composed by, from left to right, input, image, convolutional, fully connected, and output layers.

I(n)	I(n-1)	I(n-2)
Q(n)	Q(n-1)	Q(n-2)
P	T	P
|x(i)|	|x(i-1)|	|x(i-2)|
|x(i)|^2	|x(i-1)|^2	|x(i-2)|^2

(a)

I(n+1)	I(n)	I(n-1)
Q(n+1)	Q(n)	Q(n-1)
P	Y_T	Y_{TP}
|x(i+1)|	|x(i)|	|x(i-1)|
|x(i+1)|^2	|x(i)|^2	|x(i-1)|^2

(b)

Fig. 2. Image formatting: (a) PTI format and (b) proposed new format.

size and complexity. By using several filters, the convolutional layer has multiple ways of interpreting the input, keeping the network small while still being able to capture the complex nonlinear behavior of the PA. The filters inside the convolutional layer are of size 3×2.

Comparing to the prior art, the proposed model significantly improves the previously proposed [5] RVTDCNN-based architecture. Innovations, such as preprocessing and reformatting of the image layer, result in improved performance, as measured by the NMSE and ACLR, without increasing network complexity. As can be seen in Fig. 2, the image layer size remains the same (5×3). The results are improved by enriching the input data, providing a more informed basis for the network's analysis.

The rest of the network is composed of a shallow NN composed of one input layer with N neurons and one output layer. This shallow NN will take the feature maps as input and convert these features into the generated signal sample of the BM.

B. Applied Innovations to the NN

The network described in [5] was improved in two ways: on the input layer, preprocessing is applied to the data being fed to the network and the image layer format has been rearranged.

1) Input Layer—Temperature/Output Emphasized Correlation: Using the temperature parameter as input to the image layer without any preprocessing did not deliver acceptable network results. To emphasize the relationship between the temperature and the output of the PA, (1) was used to obtain $Y_T(T)$. Because the output power depends on temperature as well as the IPBO level, (1) combines the two

$$Y_T(T, \text{IPBO}) = \frac{\text{diff}(\text{mean}(\text{outputSignal}))}{\text{diff}(\text{temperatureRange})} * T. \quad (1)$$

2) Input Layer—Function of Power and Temperature: A minor change that brought considerable improvements as well was to introduce a new parameter in the form of a function between temperature and power. Instead of using power without preprocessing as the third input into the matrix (as shown in Fig. 2(a), third row and third column), it was concluded that Y_{TP} would deliver better results, by combining both temperature and power into the same parameter. The expression for Y_{TP} is presented in (2), where k was swept from 1 to 10 to find the best value, having set a final $k = 2$

$$Y_{TP}(T) = \text{mean}(Y_T, P) \wedge k. \quad (2)$$

3) Image Layer—Centering of the Samples: Centering the present sample and applying both memory taps to the right and left, with lead and a lag memory tap in the same matrix, as presented in Fig. 2(b), Fig. 2(a) shows the original PTI image layer format. This allows for both lead and lag memory taps and for the temperature and the present samples to be placed in a central position inside the image layer such that they are included in the maximum number of convolution operations, following the same principle as used in [5].

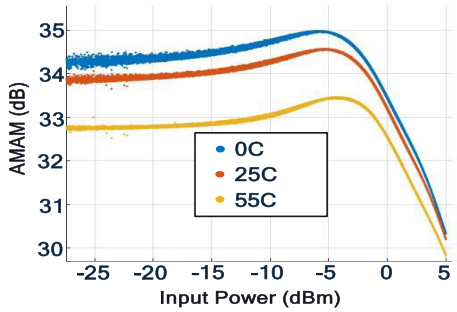


Fig. 3. PA behavior for different temperatures.

III. Results

A. Dataset Generation

The dataset was generated using pathwave advanced design system (ADS) [9], which involved using fast circuit envelope level 3 (FCE3) models for each temperature, totaling 16 models ranging from -45 °C to 85 °C, and for three different levels of IPBO from 0 to 6 dB. The aim of this dataset is to provide sufficient data to both train and validate a single predictive model for the PA that can effectively represent the range of temperatures, thereby replacing the 16 ADS models.

Fig. 3 shows a binned averaged amplitude-to-amplitude (AMAM) distortion plot, comparing the behavior of the PA at three different temperatures (0 °C, 25 °C, and 85 °C). The *IQ* samples and power and temperature data are preprocessed in MATLAB and incorporated into a matrix with a size of 5×3, using two memory taps and two nonlinear orders.

Different random seeds are utilized to generate each subset of training, validation, and testing data, enabling the comparison of the proposed predictive model with the ADS models. It is worth noting that the proposed BM is being developed to replace the 16 ADS models, and thus, it will be compared against these different models in terms of accuracy and efficiency.

B. Training

The training signal is a 100-MHz orthogonal frequency-division multiplexing (OFDM) signal. Measures for the input–output pair signal are taken at different temperatures and IPBOs and used to train the network. Two different types of training were made, one for three temperatures [0, 35, 85] °C and one for five temperatures [-45, -20, 0, 35, 85] °C. The

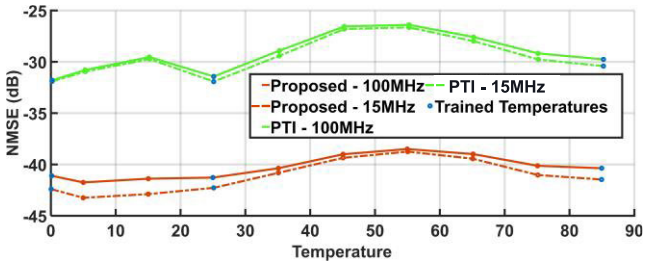


Fig. 4. NMSE values for the network's output trained for three temperatures. Taken for signals of 15 MHz in dashed line and 100 MHz in full line. The proposed network outputs are presented in orange and the PTI ones are in green.

three-temperature training supports the 0 °C–85 °C temperature range as well as 0–6-dB IPBO, and the five-temperature network covers the whole range from −45 °C to 85 °C.

The network is trained to minimize the error between the reference samples, created in ADS, and the model output samples. For this, the cost function used to train the network is the mean squared error (MSE), expressed in (3), with the reference labels being I and Q values, and the output of the network being I' and Q'

$$\text{MSE} = \frac{1}{2N}\sum_{n=1}^{N}\Big[\big(I'(n) - I(n)\big)^2 + \big(Q'(n) - Q(n)\big)^2\Big]. \quad (3)$$

The optimizer chosen was adaptive moment estimation (ADAM), with a mini-batch size of 100 samples and a maximum number of 25 epochs, with a learning rate set to 1^{-4}. These parameters were chosen based on sweeps of trainings and tests to see the best suited ones. The chosen activation function was rectified linear unit (ReLU).

The parameters referring to the number of filters in the convolutional layer and the number of neurons in the fully connected layer were chosen after empirical tests showing that the best combination would be the number of filters, $F = 10$, and the number of neurons, $N = 50$, for our dataset.

C. Testing

The testing subset was created to test the whole range of temperatures. A similar signal as the one created in the training subset with 100-MHz bandwidth was first created, as well as a 15-MHz OFDM signal. While the network was only trained for 100 MHz, it was tested for different bandwidths as well. Also, the testing subset was composed of 16 temperatures ranging from −45 °C to 85 °C to test the network generalization capabilities to the range of temperatures. In this way, the values obtained during the testing phase are compared against data generated by 16 different ADS models. This allows to directly compare the proposed approach with the ADS model performance.

After training, the testing data for the whole temperature range are fed to the model. To measure the network performance for the temperature range, the IPBO is fixed at 0 dB and the temperatures are swept from 0 °C to 85 °C for the network trained with three temperatures and from −45 °C to 85 °C for the network trained with five temperatures. The testing results from the proposed model are plotted against the ones obtained using the PTI format of the image layer and without preprocessing. Note that the network complexity is the same in both PTI and proposed cases, changing only the preprocessing and image format. In Fig. 4, the NMSE plot is traced for 100 and 15 MHz. The proposed network results are

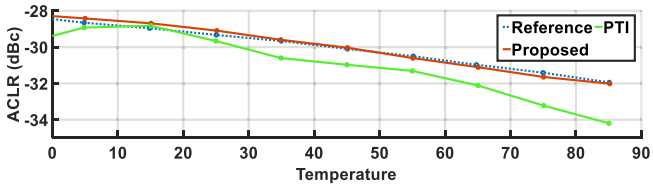
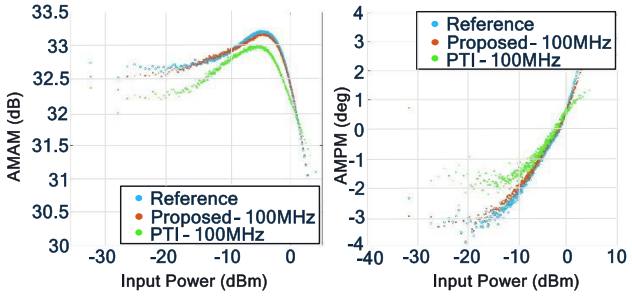


Fig. 5. ACLR values for the network's output trained for three temperatures. Taken for signals of 15 MHz in dashed line and 100 MHz in full line. The proposed network outputs are presented in orange and the PTI ones are in green.



(a) (b)

Fig. 6. (a) AMAM plot and (b) AMPM plot. Both AMAM and AMPM are taken for a signal of 100 MHz with the network trained for three temperatures.

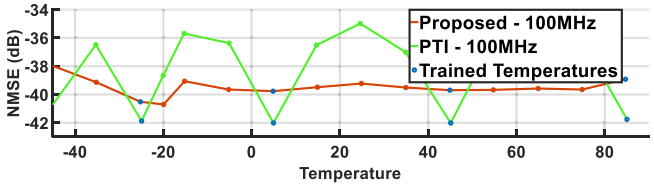


Fig. 7. NMSE values for the network output trained for five temperatures. Taken for signals of 15 MHz in dashed line and 100 MHz in full line. The proposed network outputs are presented in orange and the PTI ones are in green.

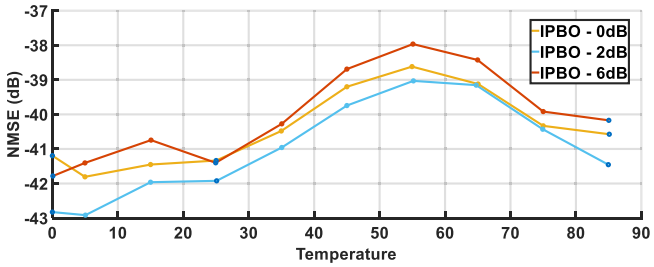


Fig. 8. NMSE values for the network output trained for three temperatures across several IPBO levels and from 0 °C to 85 °C.

shown in orange, while the PTI ones are shown in green. While the network was only trained for 100 MHz, it can perform as well for the lower BWs, presenting an average NMSE for the proposed network of −40.26 dB for 100-MHz signals and −41.01 dB for 15-MHz signals.

Fig. 5 shows the ACLR plots, again with the orange one referring to the newly proposed network, following the reference in dashed blue, while the ACLR represented in green corresponds to the PTI format. The newly proposed network shows an improvement over the PTI format.

The AMAM and AMPM plots are presented in Fig. 6, where again, it is concluded that the proposed network follows much closer the reference values when compared to the PTI network outputs.

The results for the five-temperature trained network are shown in Fig. 7, where different results are shown this time. The PTI approach (green curve) achieves better performance

in the trained temperatures, and however, it does not accurately cover the entire range of temperatures. The NMSE from the proposed approach (orange curve) achieves consistent values around the whole range of temperatures from −45 °C to 85 °C. In this case, the mean NMSE for the PTI network format is −36.48 dB, while the mean NMSE for the proposed model is −39.57 dB.

Finally, in Fig. 8, the network trained with three temperatures is tested across all the temperatures for different IPBOs. The network presents a stable PA model across the different power levels of 0-, 2-, and 6-dB IPBO. The average NMSE for the different IPBOs is −40.26, −41.03, and −40.06 dB for an IPBO of 0, 2, and 6 dB, respectively.

IV. Conclusion

In this work, an improved temperature- and power-dependent PA CNN-based model is proposed. This model includes a novel image layer format that uses preprocessed PA temperature and power parameters. The input data are arranged such that the inclusion of the most important parameters in the feature maps is promoted.

This method can work in a wide range of different PA states, tested from −45 °C to 85 °C and 0 to 6 dB. By using preprocessing of the parameters, it better compensates the effects of temperature and power variation on the PA behavior when compared to the previous method. It outperforms the previous method on equal conditions by over 3 dB when trained for five temperatures and by a larger margin when both are only trained for three temperatures.

References

[1] A. Abdelhafiz, L. Behjat, and F. M. Ghannouchi, "Generalized memory polynomial model dimension selection using particle swarm optimization," *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 2, pp. 96–98, Feb. 2018, doi: 10.1109/LMWC.2017.2783847.

[2] F. Mkadem, A. Islam, and S. Boumaiza, "Multi-band complexity-reduced generalized-memory-polynomial power-amplifier digital predistortion," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 6, pp. 1763–1774, Jun. 2016, doi: 10.1109/TMTT.2016.2561279.

[3] T. Ota et al., "A novel multi-band look-up table based digital predistorter with a single common feedback loop," in *Proc. Asia–Pacific Microw. Conf. (APMC)*, Nov. 2018, pp. 551–553, doi: 10.23919/APMC.2018.8617473.

[4] L. Sun et al., "A low complexity LUT-based digital predistortion block with new pruning method," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 9, pp. 1131–1134, Sep. 2022, doi: 10.1109/LMWC.2022.3163890.

[5] A. Motaqi, M. Helaoui, N. Boulejfen, W. Chen, and F. M. Ghannouchi, "Artificial intelligence-based power-temperature inclusive digital predistortion," *IEEE Trans. Ind. Electron.*, vol. 69, no. 12, pp. 13872–13880, Dec. 2022.

[6] X. Hu et al., "Convolutional neural network for behavioral modeling and predistortion of wideband power amplifiers," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 3923–3937, Aug. 2022, doi: 10.1109/TNNLS.2021.3054867.

[7] D. Wang, M. Aziz, M. Helaoui, and F. M. Ghannouchi, "Augmented real-valued time-delay neural network for compensation of distortions and impairments in wireless transmitters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 242–254, Jan. 2019, doi: 10.1109/TNNLS.2018.2838039.

[8] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019, doi: 10.1109/COMST.2019.2926625.

[9] (2022). *PathWave System Design-Keysight*. [Online]. Available: https://www.keysight.com/us/en/assets/3121-1074/technical-overviews/PathWave-System-Design.pdf