Enhancing Transistor Sizing in Analog IC Design using a Circuit-Focused Semi-Supervised Learning

Rayan Mina
Department of Electrical and
Mechanical Engineering
Saint-Joseph University of Beirut
Mar Roukoz, Lebanon
rayan.mina@usj.edu.lb

George E. Sakr
Virgil Systems
Toronto, Canada
gsakr@permion.ai

Houssam Nassif
Meta
Seattle, USA
houssamn@meta.com

Abstract—**This work investigates the application of artificial neural networks to predict transistor dimensions in analog integrated circuits using semi-supervised learning. Traditionally, circuit designers apply a time-consuming iterative approach to find the optimal dimensions of transistors that satisfy a set of performance metrics. To address this problem, we propose to use artificial neural networks combined with an innovative approach wherein each transistor's dimensions were first predicted using its own network to identify potential learning behavior differences. Some transistors exhibited favorable validation loss levels, while others displayed up to 3 times higher loss values. Building upon this observation, a focused approach was developed, involving the splitting of prediction tasks into two individual networks. The first targeted transistors with low validation losses and relied solely on circuit performance metrics as inputs. The second, designed for challenging transistors, introduced a novel input structure encompassing not only the performance metrics but also the dimensions of other well-trained transistors. This adjustment led to a notable reduction in both training and validation losses by 3.5 times, thus enhancing prediction accuracy for the challenging transistors. These findings underscore the importance of tailored artificial neural networks in enabling more efficient transistor sizing and present a promising approach for advancing analog integrated circuit design automation. Furthermore, this study contributes to the understanding of machine learning efficiency in the context of analog design.**

Keywords— *Transistor sizing, Analog IC design, Machine Learning, Artificial Neural Networks.*

I. INTRODUCTION

The design of analog integrated circuits (IC) is widely considered a challenging and time-consuming task since their performance depends heavily on the physical dimensions (width and length) of circuit components, namely transistors and passives. Analog designers spend considerable time tuning the design parameters DP (W_k, L_k) of all components to achieve a set of performance metrics (PM) using electrical simulators. For example, an amplifier with 5 transistors has 10 design parameters and its overall performance can be measured by several metrics such as voltage gain, thermal noise floor, common-mode rejection ratio, power supply rejection ratios, gain-bandwidth product, phase margin and current consumption. Each metric can be estimated using a proper type of simulation in the design environment. This iterative design phase is often referred to as Transistor Sizing.

An important effort has been conducted in the past decade to use artificial neural networks (ANN) to assist in this process [1] relying mainly on the efficiency of ANNs to solve complex multivariate problems. Both reinforcement and supervised learning approaches were used and tested over different circuits and technology nodes [1]. Promising results were obtained using reinforcement learning [2][3][4], however this approach is not considered in this work. Although good results were obtained in supervised learning, they were often limited to a certain region of the design space [5][6], or linked to an available set of pre-designed circuits [7][8], or to scale an already existing solution in a novel technology node [9] or to circuits with a restricted number of PM [10].

In this paper, we investigate the efficiency of the supervised learning approach in achieving accurate prediction results during component sizing phase. We show that attempts to predict the dimensions of all components using a single ANN (joint prediction) is not optimal and that the networks can learn the design parameters of some transistors more efficiently than others. Moreover, we devise an innovative approach that splits the learning into two ANNs: the first network groups the components having low validation losses with the performance metrics PM as sole inputs, while the second predicts more challenging transistors and is fed the design parameters of the first network. This allows us to gain neural insights into the task at hand [11][12]. A final contribution of this work is to explain from an electronic circuit design perspective why such learning behavior differences exist between circuit components.

This paper is organized as follows: The definition of the problem to be tackled is presented in Section II including the inputs, outputs, and the circuit under study. The joint prediction approach using a single ANN for all components is explained in Section 3 with all the relevant machine learning details. Section 4 describes the innovative study performed to identify the learning behavior differences between the circuit components. Section 5 presents how the focused approach was derived from the results of Section 4 to split the prediction into two individual neural networks having different inputs.

II. PROBLEM DEFINITION

The problem we are trying to solve in this work can be described as follows: "considering an operational amplifier (op-amp) circuit, what would be the optimal physical dimensions of its transistors (W_k, L_k) and passives (R_k, C_k) to meet a set of performance metrics?". To answer this, we will use ANN to create a model that can approximate the electrical behavior of the circuit by linking design parameters (DP) to PM based on a labeled and limited-size dataset (training phase). Also, this model should be capable of generalizing to predict new outputs for previously unseen inputs with sufficient accuracy (validation phase). In such scheme, PM are the ANN input features while DP are the outputs (Fig. 1).
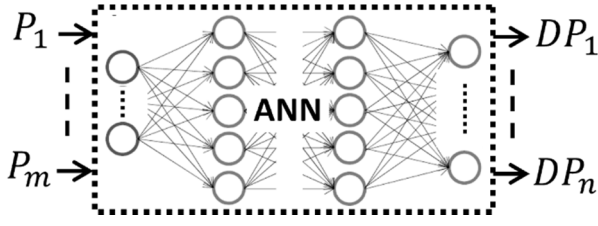
Fig. 1. ANN using the circuit's performance metrics as input features and the design parameters as outputs.

Several relevant questions arise to solve the problem:

- What do we define and obtain performance metrics?

- How to generate the data used in the training phase?

- How to generate the data used in the validation phase?

- What is the adopted architecture of the ANN?

In the following paragraphs we address these questions We will first describe the studied circuit, then how the dataset and the subsets were generated and finally the details of the adopted machine learning approach.

A. The Studied Circuit

There are various analog circuits designed today for numerous applications such as signal amplification, signal filtering, frequency mixing and others. A well-known circuit that is extensively used in closed-loop systems is the op-amp. Therefore, in this work we have decided to study a basic CMOS op-amp with resistive biasing and having a source-follower output stage. The circuit schematic is given in Fig. 2.

The op-amp example is composed of 8 transistors in total, of which six share the same design parameters two by two (W_1,L_1 / W_2,L_2 / W_3,L_3). Therefore, there are 5 distinct transistors only, each having a width and a length, and one resistor value resulting in 11 design parameters. The op-amp overall performance can be evaluated using a set of metrics, that need to be defined properly.

In this work, we have chosen to use the metrics summarized in TABLE I. , which are well-known to the analog IC design community as well as in the op-amp standard theory. They capture the overall circuit behavior in terms of dc, ac, frequency, noise, and nonlinearity.
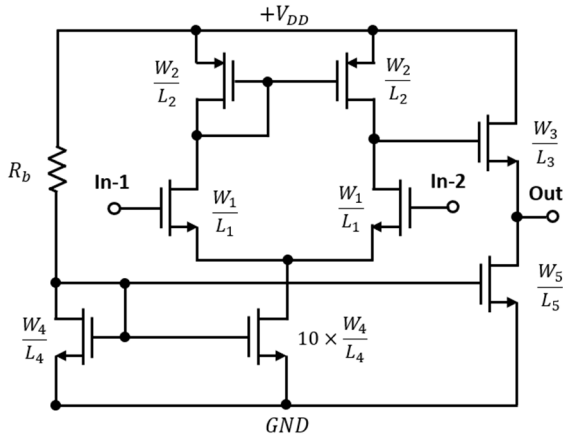

Fig. 2. Simple CMOS op-amp composed of an amplifying stage and a driving source-follower output stage for low resistive loads.

TABLE I.　　PERFORMANCE METRICS OF THE STUDIED CIRCUIT

Symbol	Performance Metrics PM	
	Description	*Note*
A_V	Open-loop voltage Gain	
CMRR	Common-mode rejection ratio	Unwanted signals that are common to inputs
PSRR+	Positive supply rejection ratio	Noise coming from positive power lines
NSRR-	Negative supply rejection ratio	Noise coming from negative power lines
IIN	Integrated noise level	
NTH	Thermal noise floor	
F_C	Noise corner frequency	
GBW	Gain-bandwidth product	
PhM	Phase margin	
Ibs	Current consumption	
V_{OUT}	dc output level	
IIP_3	3rd order Input-intercept point	

B. Dataset Generation

The semi-supervised setting relies on an automatic generation and labelling of the training dataset. This task was performed using a SPICE-like electrical simulator. We created a python script that generates one file containing a user-defined number of DP samples. Those samples have random values and are uniformly created between predefined bounds that are consistent with the design space of the circuit and the chosen technology (CMOS 0.13μm). TABLE II. shows the limits used in our study. The file is fed as an input to the electrical simulator that automatically computes the corresponding PM values. Since this operation requires high-speed hardware, this dataset was generated in 6 hours of runtime using a 16-core i9-12900 12th generation processor. Each row of the dataset is composed of 23 columns: 11 values for the design parameters and 12 for the performance metrics. In total we generated 150,000 samples.

C. Machine-Learning Approach

We now describe the whole learning process for our ANN, and detail our machine learning approach.

1) Dataset separation: To ensure the robustness and generalization capability of the ANN models, the generated dataset was partitioned into training, validation and test sets, with a ratio of 80%, 10% and 10% respectively. The training dataset is used to learn the neural network parameters while the validation set is used to test its performance. Finally the architecture that exibits the smallest validation loss is used on the test set. This separation prevents overfitting, and generalize to new data. The computed validation loss VL, provides a quantitative measure of how well the model is likely to perform in real-world scenarios.

TABLE II.　　RANDOM DATASET GENERATION BOUNDS

Symbol	Bounds		
	Description	*Minimum*	*Maximum*
W_K	Width of MOS transistors	0.3μm	50μm
L_K	Length of MOS transistors	0.13μm	1μm
R_b	Value of the biasing resistor	5kΩ	40kΩ

2) Data filtering and loss function: In our application we are trying to predict DP from PM. To perform well in the case

of op-amp, MOS transistors need to operate far from the deep triode region, ideally in saturation. No matter how well we put an effort to bound the random generation of the dataset to guarantee this condition, there will be minor samples in our data where the transistors are poorly biased and hence will act in deep triode region. Those points are actually outliers, and will lead most likely to a poor circuit performance from which the ANN is building its model. As it is more beneficial to push the network to learn on the majority of the "stable" points, we designed a logic-based filtering function that rejects samples that fall into the deep triod region. Finally, we adopt the Mean Absolute Error (MAE) as a training loss function since it was proven to be one of the most stable loss functions for regression problems [13].

3) Activation of neurons: The primary role of activation functions in ANN is to transform the summed weighted input of a neuron into an output value to be fed to the next layer or to the output. They are mainly used to add non-linearity to the networks. Since all circuit design parameters are positive quantities (Widths, Lengths), Rectified Linear Unit (ReLu) is usually recommended and widely-used in such cases [14]. In this work, we have opted for the Exponential Linear Unit (ELU) a similar variant of ReLU that modifies the slope of the negative part of the function to allow for slow smoothing and to avoid bias shift effects [14].

4) ANN architecture: The 12 performance metrics of the op-amp comes from various types of electrical simulations: dc, ac, frequency, noise, and large-signal. Moreover, MOS transistors have non-linear behavior that considerably affects the op-amp performance. The number and the diversity of the simulations used, make the relationships between DP and PM excessively difficult to model with moderate mathematical models. Due to this complexity and to the non-linear behavior of the op-amp circuit, a fully-connected flat ANN architecture is devised in this work.

5) Training Process: The network is trained multiple times with different random initial weights for the neurons (iterations) and different random splits between the training and validation subsets (random states). The training is implemented in a developed Python code based on TensorFlow using Keras library and Adam optimizer.

III. JOINT APPROACH: PREDICTION OF ALL TRANSISTORS SIMULTANEOUSLY

Our first attempt to train the ANN using the previously generated dataset is based on the straightforward method described by Fig. 1: all PM are used as inputs and the same network will be trained to predict all the DP simultaneously. We call it the joint approach since all the design parameters are jointly predicted by the same ANN. This network, which has 12 inputs and 11 outputs, will use 120,000 samples (80%) within the dataset to learn the links or relationships between PM and DP. The training in the joint approach tries to build approximations of the functions $DP_{k = 1...11} = f(P_1, ..., P_{12})$ blindly as a black box, without any prior imposed condition or any guidance from a human designer. We trained on different network complexities in terms of number of layers (input, hidden and output) and number of neurons per layer, to look for the best architecture that provides the lowest validation loss (VL). TABLE III. provides details on the different ANN complexities we have used in this work. Each result is

obtained by repeating the process for 5 different random states and 5 different iterations (25 points). Each single point required 9 minutes of processing using RTX 3090 Ti GPU leading to ~4 hours of runtime for each result that corresponds to just one network complexity. By doing a progressive training process our model has converged, and losses have stabilized after 600 epochs of training for all network configurations. Fig. 3 shows an example of training results for the case of a flat network having 7 layers and 150 neurons per layer. It is clear how training loss and VL have reached a plateau and that no more processing is needed.

We trained 24 different networks as per TABLE III. guidelines and summarize the results in Fig. 4 by taking the average loss values from the 25 points of each result.

TABLE III. ANN COMPLEXITIES USED IN THE JOINT APPROACH

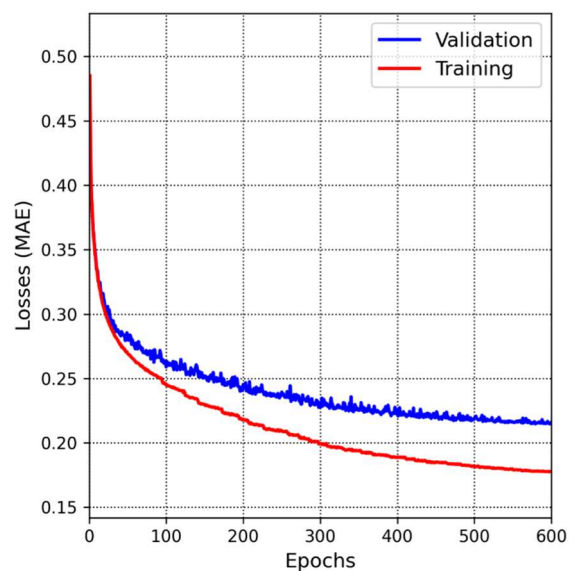| Symbol	ANN Complexities			
	Description	*Minimum*	*Maximum*	*Step*
L	# of layers	3	8	1
N	# of neurons per layer	50	200	50

Fig. 3. Training and Validation losses on one ANN case showing how both losses have stabilized after 600 epochs.

A couple of observations can be made from Fig. 4: First, with more network complexity (layers, neurons) the training loss decreases since the ANN is more capable of learning the complex and non-linear circuit relationships between PM and DP. Second, although the validation loss improves when we add more layers to the ANN, the decrease rate is slower, and it behaves differently since it reaches a minimum range between 0.22 and 0.25 no matter how many layers or neurons we add. Worth noting that increasing the network complexity clearly does not help, as we can see how the magenta, green and yellow curves at the right plot of Fig. 4 start to increase indicating a poor generalization capability of the learning. The MAE validation loss cannot be reduced to less than 25% on the 11 design parameters. In other words, the width and length of the transistors as well as the resistor value may exhibit an error of 25% or higher during ANN prediction. Of course, this error is averaged on the eleven DP, therefore some DP may have a higher error while other ones may have a lower error.
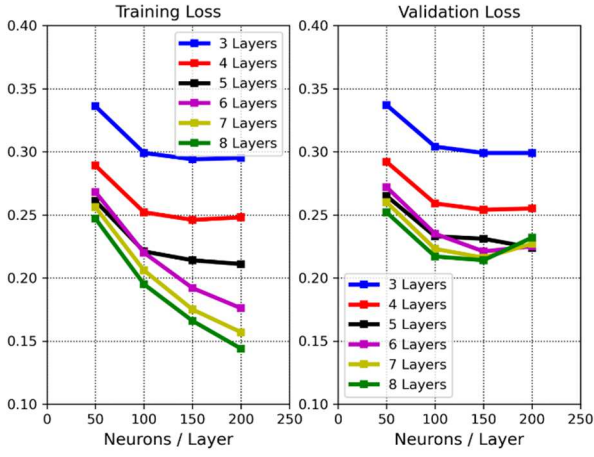
Fig. 4. Training and Validation losses for all ANN complexities showing the limitation of the learning efficiency in the joint approach.

Our results suggest that there is an inherent minimum on VL obtained in the joint approach that represents a limitation of the model accuracy. In the next section we will devise an alternative ANN learning strategy to better understand the reasons and overcome this effect.

IV. INVESTIGATIVE APPROACH: PREDICTING TRANSISTORS INDEPENDENTLY

To overcome the limitation observed in the joint approach, the first step is to understand why we reached a minimum for VL. This section describes an innovative domain-augmented approach that links circuit and ANN learning behaviors.

A. Description of our domain-augmented approach

An IC engineer has a domain knowledge of the circuit behavior and can expect in advance what transistor(s) or component(s) influence a certain performance metric. During simulations, circuit sizing never happens in one shot [15][16], it is an iterative process, where each component is tuned as a standalone entity. For example, an IC designer will not attempt to size the width of one transistor without sizing its length. Moreover, the focus on a certain component should be mapped to the type of simulations being run [17][18][19], e.g., during op-amp ac voltage gain simulation the focus is mainly on the differential pair transistors (M_1 and M_2) as they directly impact this specific performance. Therefore, the core idea in this section is to assign a different ANN for each transistor to predict its width and length simultaneously, as any IC designer would do in real life. The benefit of this study is to capture whether the behavior of some transistors can be more easily modeled, and in parallel if some are more challenging. The bloc diagram of this innovative approach is depicted in Fig. 5. Each circuit component is treated as a standalone entity and is assigned a network to predict its dimensions. In this work, we ended up having 6 individual ANNs to model the circuit (each fed the 12 metrics PM) compared to just one ANN in the joint approach. This process will naturally increase the computation time and complexity of the system. Nevertheless, we are not planning to predict the dimensions this way, this separation will be used only to understand the tricky relationships between DP and PM.

B. Training Results

To make sure the results are stable and can be used to draw conclusions, we have tested our investigative approach on six different networks having 4, 6 and 8 layers with 50 and 100

neurons per layer for each case. The values obtained for training loss and VL are given in Fig. 6.
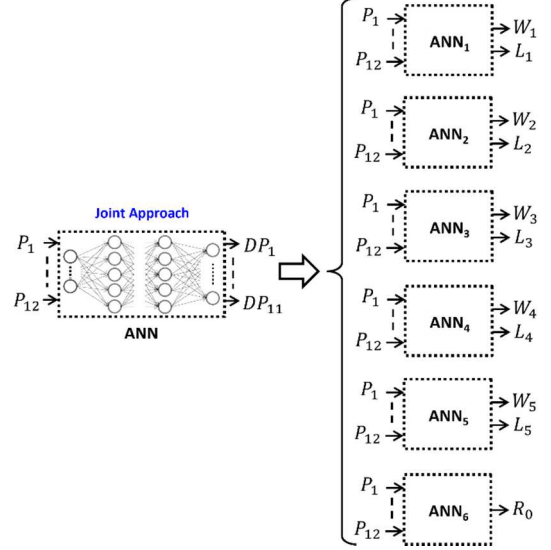

Fig. 5. Dividing the learning process by mapping an individual ANN for each circuit component (Transistor, resistor) in the investigative approach.
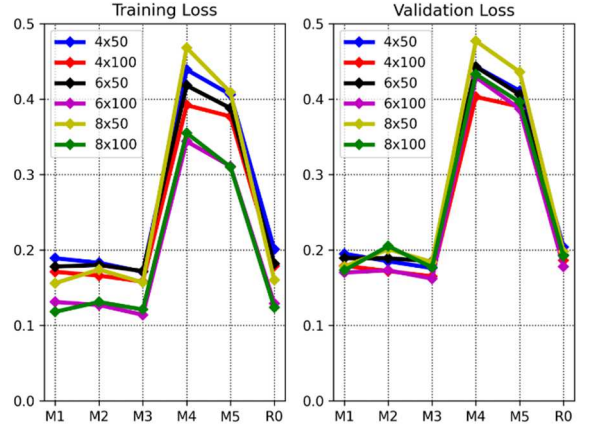

Fig. 6. Validation and Training Losses showing that the circuit components are not similarly learned by ANN. Two transistors, M_4 and M_5, show significant high VL while the others have much lower values.

C. Physical Interpretation of the Results

It is clear from the two peaks in all curves of Fig. 6 that transistors M_4 and M_5 are challenging components for the ANN to learn. In fact, they both exhibit 3 times higher training and validation losses compared to the remaining transistors M_1, M_2, M_3 and the resistor R_0. This observation means the relationships between W_4 / L_4, W_5 / L_5 and the performance metrics PM are less meaningful than those between the remaining design parameters and PM. Most likely M_4 and M_5 are responsible for the minimum bound on VL observed using the joint approach in the previous section, when one single ANN was trying to predict all DP simultaneously.

To understand the previous anomaly on the challenging transistors one must refer to a circuit designer's perspective. M_4 transistors (left and right on schematic of Fig. 2) serve as a current mirror to fix the dc current flowing in the symmetrical branches of the differential pair M_1 and its load M_2. Their role in the circuit is indeed important to properly bias the active pair of the op-amp. However, M_4 has limited effect on the performance metrics described in TABLE I. . In fact, any IC designer with sufficient skills in the field would

know that M_4 impacts mainly CMRR and NSRR due to the dependency of those quantities on its output resistance r_o and has little effect on the remaining PM. Similarly, M_5 acts as a load to the active transistor of the source follower M_3. Its role is similar to a simple resistor and will affect only the second stage voltage gain which is very low (almost unity ~1). Therefore, sub-optimal values for W_5 / L_5 will certainly make A_V drop, but once they are in a typical range their effect is almost negligible on the remaining PM.

Based on the VL results obtained during ANN learning, we have identified two patterns: circuit components with low VL values and others with significantly high values. Moreover, the previous physical interpretations from a circuit design perspective corroborate those findings. Building upon this observation, we will propose in the next section a focused approach to make the learning more efficient and reduce the VL value below the minimum obtained in the joint approach.

V. FOCUSED APPROACH: SPLITTING OF NEURAL NETWORKS

The core idea in the focused approach is to split the prediction of the design parameters into two individual ANNs based on the VL patterns obtained from section IV.
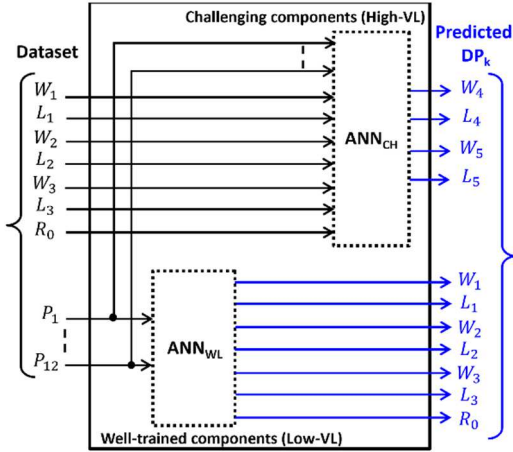

Fig. 7. The focused approach wherein prediction is split into 2 distinct ANNs based on VL patterns. ANN_{CH} predicting the challenging transistors incorporates design parameters of other transistors as inputs.

Transistors M_1, M_2, M_3 and resistor R_0 showed low VL values and will be hence predicted together using a single ANN that learns from the performance metrics solely (ANN_{WL} in Fig. 7). On the other hand, a second network will predict the challenging transistors M_4 and M_5 and has a distinctive and unique input structure incorporating (in addition to PM) dimensions of the other transistors from the first network (ANN_{CH} in Fig. 7). Details on this split are shown in Fig. 7. The focused approach is guided by the VL results obtained in section IV, but it is also motivated by the physical interpretation of the roles of M_4 and M_5 in the circuit. We strongly believe that the ANN predicting them had difficulties capturing the relationships between their DP and PM because there is little link existing between those quantities. By adding the design parameters of the other components with much lower VL values to the ANN input, the latter has now more relevant information to process and can couple information from DP of M_1, M_2, M_3 and R_0 with the performance metrics.

We have tested the focused approach by training 12 different networks (4,6,8 layers and 50,100,150,200 neurons per layer), we summarize the results in TABLE IV. . We repeated each result 25 times using the same process explained

in paragraph II.C.5). In TABLE IV. , both losses have decreased significantly by a factor of ~ 3.5 (30% of its original value) when comparing the focused approach results with the single transistor results of section IV. For example, in a 6-layer network having 100 neurons per layer, the VL decreased from ~0.4 down to ~0.11. This decrease is significant and can considerably enhance the prediction accuracy for the challenging transistors, which are the main issue faced in the sizing problem. For the remaining circuit components, transistors M_1, M_2, M_3 and resistor R_0, the results were quasi-equal to those in section IV. This is expected since no additional inputs were given to ANN_{WL} that predicts those components (see Fig. 7).

Based on the training results obtained on ANN_{CH}, there is a clear enhancement in the training loss, and more importantly in VL, hence in the prediction accuracy. To measure this improvement, we compare the results of the focused approach with the joint approach if section III. Since we have two networks and two distinct values for VL (one for ANN_{WL} and one for ANN_{CH}), we compare the average loss. As Fig. 8 shows the VL decreased by ~50% on all 12 networks tested, a considerable improvement by the focused approach leading to better prediction accuracy. TABLE IV. provides also the test loss values (TL) computed on the last 10% test set of the data.

TABLE IV.   LOSS IMPROVEMENT FOR CHALLENGING TRANSISTORS

ANN Architecture		Investigative Approach (ANN for each transistor)		Focused Approach	
		VL	VL	VL	TL
Layers	Neurons per Layer	M4	M5		
4	50	0.44	0.41	0.16	0.18
4	100	0.40	0.39	0.12	0.13
6	50	0.44	0.41	0.15	0.16
6	100	0.43	0.39	0.11	0.12
8	50	0.48	0.44	0.13	0.13
8	100	0.43	0.40	0.10	0.11

A major outcome of this work is highlighting that for some circuit components, the joint approach which learns to predict all the DP together, has serious limitations. Putting all circuit components in the same ANN and disregarding their different roles within the circuit leads to sub-optimal results. Splitting the process into two distinct networks based on VL patterns obtained from the training of each circuit component alone enables an efficient and circuit-focused learning approach. Results in Fig. 8 for 12 different ANNs confirm this finding.
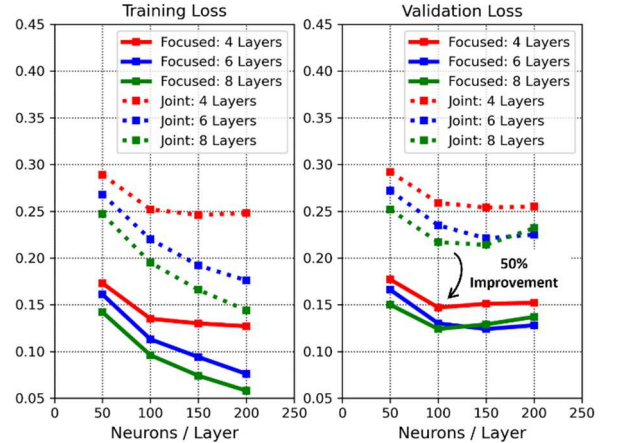

Fig. 8. Comparison of the average VL in the focused approach with the VL in the joint approach, showing 50% improvement.

Building upon this, we recommend a general 3-step design procedure for analog circuit sizing based on semi-supervised learning and a hierarchical ANN training process:

- Evaluate the difficulty of predicting each circuit component alone by training an ANN for each component using PM as sole inputs.

- Classify components into two categories, challenging ones exhibiting high VL values and straightforward ones with low VL values.

- Split the learning process into a network for each category, with the second ANN having a modified input structure incorporating DP of other components.

From a circuit design perspective, the challenging transistors of the op-amp circuit studied in this work have less impact on the PM compared to the straightforward ones. First, M_1 and M_2, namely the active pair and load are the central components that make the main circuit function: the amplification of the inputs. Those two components have a major impact on almost all PM since the input signals are fed and amplified in this part of the circuit. M_3 also has an important role in the second stage of the circuit since it is the active component of the source follower connecting the amplified signal of the first stage to the output where all PM are measured. Therefore, it is reasonable to expect an ANN to learn the behavior of those components from the PM solely. On the other hand, M_4 is used to copy a current fixed by resistor R_0 and convey it to the differential pair branches. The dimensions of M_4 are not particularly critical to this function since M_4 left transistor is diode-connected and will be in saturation regardless of the values chosen for W_4 and L_4. The latter have impact on few PM (mainly CMRR, NSRR and to a certain extent I_{bs}) only. However, they are affected by the behavior of M_1 and M_2 as the sum of drain-source voltage of M_1, M_2 and M_4 is equal to the supply voltage V_{DD} (Kirchoff law). In addition, M_5 acts a simple load to M_3, therefore it does not need to operate in saturation to perform well and W_5/L_5 also have little impact on most PM. Similarly, they are affected by the behavior of M_3.

VI. Conclusion

This paper investigated several approaches in using supervised learning and ANNs to predict components dimensions from analog circuit's performance metrics. It shows that joint prediction of all components of an op-amp using one network led to significant challenges and sub-optimal results with a high validation loss. To overcome this, we conducted an investigative approach to classify transistors based on high (challenging) and low VL patterns obtained from the training of a single ANN for each circuit component. Then, we split the learning into two distinct networks one for the challenging transistors that incorporates the design parameters from the remaining transistors as additional inputs, and the second network is fed PM as sole inputs. Results show considerable improvement with VL values decreasing to 50% compared to the joint approach values. We provide a physical interpretation of the split from a circuit design perspective based on the role of the circuit components. An aspect for future work is to test and validate this approach on more analog circuits, and also to integrate this approach into an analog design environment software.

References

[1] R. Mina, C. Jabbour, and G.E. Sakr, "A review of machine learning techniques in analog integrated circuit design automation" Electronics, vol. 11, issue 3, 435, 2022.

[2] H. Wang, J. Yang, H.-S. Lee, S. Han, "Learning to Design Circuits" arXiv 2018, arXiv:abs/1812.02734.

[3] K. Settaluri, A. Haj-Ali, Q. Huang, K. Hakamaneshi, B. Nikolic, "AutoCkt: Deep Reinforcement Learning of Analog Circuit Designs" In Proceedings of the 2020 Design, Automation Test in Europe Conf. Exhibition (DATE), France, 2020, pp. 490–495.

[4] Z. Zhao, L. Zhang, "Deep Reinforcement Learning for Analog Circuit Sizing" In Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, 2020, pp. 1–5.

[5] M. Fukuda, T. Ishii, and N. Takai, "OP-AMP sizing by inference of component values using machine learning" In Proceedings of the 2017 Inter. Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Xiamen, China, 2017, pp. 622–627.

[6] Z. Wang, X. Luo, "Application of Deep Learning in Analog Circuit Sizing" In Proc. of the 2nd Inter. Conference on Computer Science and Artificial Intelligence, Shenzhen, China, 2018, pp. 571–575.

[7] N. Lourenço, et al., "On the Exploration of Promising Analog IC Designs via Artificial Neural Networks" In Proceedings of the 2018 15th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), Prague, Czech Republic, 2–5 July 2018, pp. 133–136.

[8] M. Harsha, B.P. Harish, "Artificial Neural Network Model for Design Optimization of 2-stage Op-amp" In Proceedings of the 2020 24th International Symposium on VLSI Design and Test (VDAT), Bhubaneswar, India, 23–25 July 2020; pp. 1–5.

[9] R.A. Vural, N. Kahraman, B. Erkmen, and T. Yildirim, "Process independent automated sizing methodology for current steering DAC" Inter. Journal of Electronics, vol. 102, issue 10, 2015, pp. 1713–1734.

[10] S. D. Murphy, K. G. McCarthy, "Automated Design of CMOS Operational Amplifier Using a Neural Network" In Proceedings of the 2021 32nd Irish Signals and Systems Conference (ISSC), Athlone, Ireland, 10–11 June 2021; pp. 1–6.

[11] F. Kong, et al., "Neural Insights for Digital Marketing Content Design" In Proceedings of the Inter. Conf. on Knowledge Discovery and Data Mining (KDD), Long Beach, USA, 6-10 August 2023, pp. 4320-4332.

[12] S. Geng, H. Nassif, and C.A. Manzanares, "A Data-Driven State Aggregation Approach for Dynamic Discrete Choice Models" In Proceedings of the 2023 Conference on Uncertainty in Artificial Intelligence (UAI), Pittsburgh, USA, 2023, pp. 647-657.

[13] J. Qi, et al., "On Mean Absolute Error for Deep Neural Network Based Vector-to-Vector Regression" in IEEE Signal Processing Letters, vol. 27, 2020, pp. 1485-1489.

[14] D. Kim, J. Kim, and J. kim, "Elastic exponential linear units for convolutional neural networks" Neurocomputing, vol. 406, issue 10, 2020, pp. 253–266

[15] E. Rémond, E. Nercessian, C. Bernicot, R. Mina, Mathematical approach based on a "Design of Experiment" to simulate process variations. In Proceedings of the 2011 Design, Automation and Test in Europe, Grenoble, France, 14–18 March 2011; pp. 1–5.

[16] F. Montaudon, et al., A Scalable 2.4-to-2.7GHz Wi-Fi/WiMAX Discrete-Time Receiver in 65nm CMOS. In Proceedings of the 2008 IEEE International Solid-State Circuits Conference—Digest of Technical Papers, San Francisco, CA, USA, 3–7 February 2008; pp. 362–619.

[17] L. Joet, et al., Advanced 'Fs/2' Discrete-Time GSM Receiver in 90-nm CMOS. In Proceedings of the 2006 IEEE Asian Solid-State Circuits Conference, Hangzhou, China, 13–15 November 2006; pp. 371–374.

[18] M. T. Nguyen, et al, Direct delta-sigma receiver: Analysis, modelization and simulation. In Proceedings of the 2013 IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, 19–23 May 2013; pp. 1035–1038.

[19] G. Wagner, R. Mina, Amplifier for a Wireless Receiver, 2015, US patent 9,077,302. patents.google.com/patent/US20140035675A1/en.