RESEARCH Open Access

The application of deep learning technology in integrated circuit design

Lihua Dai^{1*}, Ben Wang², Xuemin Cheng¹, Qin Wang¹ and Xinsen Ni¹

*Correspondence: Lihua Dai dai_lihua@outlook.com 1 School of Integrated Circuits and Communications, Suzhou Vocational Institute of Industrial Technology, Suzhou, Jiangsu 215104, China 2 Microsoft (China) Co., LTD. Suzhou Branch, Suzhou, Jiangsu 215000, China

Abstract

This study addresses the intricate challenge of circuit layout optimization central to integrated circuit (IC) design, where the primary goals involve attaining an optimal balance among power consumption, performance metrics, and chip area (collectively known as PPA optimization). The complexity of this task, evolving into a multidimensional problem under multiple constraints, necessitates the exploration of advanced methodologies. In response to these challenges, our research introduces deep learning technology as an innovative strategy to revolutionize circuit layout optimization. Specifically, we employ Convolutional Neural Networks (CNNs) in developing an optimized layout strategy, a performance prediction model, and a system for fault detection and real-time monitoring. These methodologies leverage the capacity of deep learning models to learn from high-dimensional data representations and handle multiple constraints effectively. Extensive case studies and rigorous experimental validations demonstrate the efficacy of our proposed deep learning-driven approaches. The results highlight significant enhancements in optimization efficiency, with an average power consumption reduction of 120% and latency decrease by 1.5%. Furthermore, the predictive capabilities are markedly improved, evidenced by a reduction in the average absolute error for power predictions to 3%. Comparative analyses conclusively illustrate the superiority of deep learning methodologies over conventional techniques across several dimensions. Our findings underscore the potential of deep learning in achieving higher accuracy in predictions, demonstrating stronger generalization abilities, facilitating superior design quality, and ultimately enhancing user satisfaction. These advancements not only validate the applicability of deep learning in IC design optimization but also pave the way for future advancements in addressing the multidimensional challenges inherent to circuit layout optimization.

Keywords Integrated circuit, Circuit design, Deep learning

Introduction

With the rapid development of information technology, integrated circuits (IC) have become indispensable core components of modern electronic equipment, widely used in communication, computing, consumer electronics, medical, military and other fields. As Moore's Law approaches its physical limits, IC design faces unprecedented challenges:

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit [http://creativecommons.org/licenses/by/4.0/.](http://creativecommons.org/licenses/by/4.0/)

on the one hand, device size continues to shrink, making design complexity grow exponentially, and the risk of design errors and manufacturing defects increases; on the other hand, the market demand for high performance, low power consumption, low cost, and fast time-to-market is increasingly urgent (Afacan et al. [2021\)](#page-19-0). Integrated circuit design is complex, and its design flow is shown in Fig. [1](#page-1-0).

In this context, deep learning, as a revolutionary technology in the field of artificial intelligence, is gradually being explored and applied to integrated circuit design due to its powerful pattern recognition, feature extraction and data processing capabilities. Deep learning can learn complex nonlinear relationships in integrated circuit design through large-scale data training, so as to realize automatic design decision-making, performance prediction and fault detection tasks, which is expected to greatly improve design efficiency, shorten design cycle, reduce design cost, and promote design quality improvement (Barnwal and Dhawan [2020\)](#page-19-1). In recent years, scholars at home and abroad have carried out extensive and in-depth research on the application of deep learning in integrated circuit design. In terms of circuit layout optimization, researchers use convolutional neural networks (CNNs) to extract and classify features from layout graphs,

Fig. 1 Flow chart of integrated circuit design

and propose a series of optimization strategies that significantly improve layout quality and design convergence speed. In addition, for performance prediction, especially for power consumption and delay prediction, recurrent neural networks (RNNs) and long short-term memory networks (LSTM) are used to model circuit behavior because they can effectively process sequential data, and more accurate prediction models are realized. In the field of fault detection and diagnosis, anomaly detection algorithms based on generative Adversarial networks (GANs) and models combined with transfer learning show great potential in identifying potential design flaws. Internationally, a number of top semiconductor companies, including IBM, Intel and TSMC, have invested a lot of resources in relevant research and have achieved preliminary results (Bogaerts et al. [2019](#page-19-2)).

For example, special attention should be paid to the use of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for chip feature extraction and pattern recognition, which show great potential for improving design efficiency and accuracy (Chong et al. [2018\)](#page-19-3). At the same time, innovative applications integrating attention mechanisms and generative adversarial networks (GANs) provide a new perspective for solving complex design space exploration problems. In addition, the exploration of automatic machine learning (AutoML) in IC design parameter optimization is also the key to strengthen the research of modernity. Further, the research should also cover the application of deep reinforcement learning (DRL) in integrated circuit design automation, which has demonstrated excellent capabilities for large-scale design space search, layout optimization, and resource allocation problems (Chovan and Uherek [2018\)](#page-19-4). DRL continuously learns the optimal strategy by simulating the interaction process with the environment, which greatly improves the quality and efficiency of design decisions. DRL is one of the core technologies to promote the intelligent transformation of IC design. It is worth noting that with the great success of the transformer model in the field of natural language processing, it has also begun to emerge in the efficient processing and feature learning of integrated circuit design data. The self-attention mechanism of Transformer model can better capture long-distance dependencies and has potential advantages for dealing with high-dimensional feature associations in complex chip design, which is worth further discussion in references (Dwivedi et al. [2021\)](#page-19-5). In addition, in the face of increasing design complexity and shortened product development cycle, research on lightweight deep learning models is also particularly important. These models can reduce computational resource requirements while ensuring prediction accuracy, accelerate design iteration process, and have practical significance for rapid prototyping verification and cost control.

In view of the broad application prospects of deep learning in integrated circuit design and the current research progress, this research aims to comprehensively explore and deepen the research in this field, including but not limited to: (1) Deep learning models and algorithm optimization: for specific tasks in integrated circuit design, research and development of suitable deep learning models, such as improved CNN architecture for finer layout optimization, and customized RNN/LSTM models to improve the accuracy of performance prediction. (2) Large data set construction and feature engineering: Considering the particularity and scarcity of IC design data, this study will explore effective data acquisition methods, construct high-quality training data sets, and conduct feature selection and engineering to ensure effective training and generalization of models. (3) Cross-level optimization strategy: Study how to integrate deep learning technology in different stages of integrated circuit design (such as logic synthesis, layout, routing, etc.) to achieve global optimization from system level to transistor level and improve overall design performance. (4) Fault prediction and adaptive repair mechanism: use deep learning technology to predict possible circuit fault points in advance, and design corresponding adaptive repair strategies to reduce problems in the later physical verification stage and accelerate design convergence.

Through the in-depth exploration of the above research contents, this study is expected to provide a set of automatic design and optimization framework based on deep learning for the field of integrated circuit design, promote the intelligent and efficient direction of integrated circuit design, and further promote the rapid development of the entire electronic information technology.

Literature review

Integrated circuit design process overview

Integrated circuit (IC) design is a highly complex multi-step process involving numerous stages from concept to final product manufacturing. Generally, this process can be roughly divided into three main parts: front-end design (FE), back-end design (BE), and verification (Chong et al. [2018](#page-19-3)).

The front-end design mainly includes specification formulation, logic design and function verification. In the specification development stage, the design team defines the functions, performance indicators and application scenarios of the chip. The logic design phase uses hardware description languages (HDL), such as Verilog or VHDL, to write code to describe the behavior and structure of circuits. The HDL code is then transformed into a gate-level netlist by logical synthesis, and functional verification is performed at this stage to ensure that the design meets expectations (Chovan and Uherek [2018](#page-19-4)).

Back-end design focuses on layout design, physical verification, and manufacturing documentation. Layout is the core link, in which layout involves assigning the spatial location of various components of the circuit on the silicon chip, and routing determines the connection path between these components. Subsequently, physical verification checks whether the design meets criteria such as Electrical Rule Check (ERC), Design Rule Check (DRC), etc. (Bogaerts et al. [2019](#page-19-2)). Finally, the generated GDSII file is used to guide the actual fabrication of the chip.

Major challenges in design

As process nodes shrink, IC design complexity increases dramatically. The scale-up of designs leads to the problem of "design explosion", i.e. the complexity of design verification increases exponentially with circuit size. At the same time, the variables in the design increase, and the exploration of the design space becomes more difficult, requiring the support of advanced algorithms and computational resources (Chovan and Uherek [2018](#page-19-4)). Modern electronic devices have dual requirements for high performance and low power consumption. In mobile devices, data centers and other fields, energy efficiency ratio has become a key indicator. Designers need to find the right balance between increasing computing speed and reducing power consumption, which often requires innovative strategies at the circuit architecture, algorithm level, and physical

design (Dwivedi et al. [2021](#page-19-5)). With the progress of technology, uncertainties in manufacturing process (such as random doping fluctuation, line edge roughness, etc.) have increasingly significant effects on chip performance. DFM requires that these variations in the manufacturing process be considered early in the design process, and that their adverse effects be reduced through design optimization to ensure chip consistency and reliability (Errando-Herranz et al. [2020](#page-19-6)).

Analysis of the limitations of traditional design methods

Although the existing IC design process has made significant progress over the past few decades, its limitations have become increasingly evident in the face of increasing challenges. The traditional design process relies heavily on manual intervention and empirical judgment, especially in the layout and design optimization stages, which is not only time-consuming but also prone to human error. As the size of the design grows, this manual adjustment becomes less feasible (Garcia-Sciveres [2023](#page-19-7)). Design convergence refers to the process of completing a design under all performance, area, and power constraints. Due to the complex interactions among design parameters, traditional design methods are difficult to efficiently explore large design spaces, resulting in long design convergence time and high cost (Guo et al. [2020](#page-19-8)). With the rapid evolution of process technology, new physical effects and manufacturing constraints continue to emerge, and traditional design methods and tools often lag behind these changes, and it is difficult to solve new problems brought about by new technologies in a timely and effective manner (Hao et al. [2021\)](#page-19-9).

In-depth analysis of traditional method limitations and advantages of proposed deep learning approach

Traditional IC design methodologies, despite their historical contributions, struggle to cope with the escalating complexities and dimensionalities inherent in modern design spaces. Manual intervention, a cornerstone of conventional design, introduces subjective biases and inefficiencies, leading to prolonged design cycles and potential suboptimal solutions. The iterative nature of design convergence, exacerbated by the explosion in design variables, demands an immense computational burden that often exceeds the capabilities of traditional optimization algorithms (Hong et al. [2022](#page-19-10)).

In contrast, the proposed integration of deep learning offers transformative advantages. By leveraging neural networks' capacity for automated feature extraction and pattern recognition, our approach alleviates the need for exhaustive manual tuning, thereby accelerating the design optimization process (Khan et al. [2019](#page-19-11)). Deep learning models can effectively navigate vast design spaces, identifying Pareto-optimal solutions that balance power, performance, and area (PPA) considerations more comprehensively than conventional methods (Lai et al. [2022](#page-19-12)).

Moreover, the predictive prowess of these models enables early-stage estimation of critical design parameters, facilitating proactive adjustments to mitigate potential issues, such as manufacturing variability (Lambrechts et al. [2024](#page-19-13)). This capability aligns with the principles of Design for Manufacturing (DFM), ensuring design robustness and reliability from the outset.

A comparative analysis against state-of-the-art techniques underscores the superiority of our deep learning-driven framework. Unlike iterative optimization algorithms or rule-based methods, which may converge to local optima or fail to capture intricate design interactions, our deep learning models demonstrate higher accuracy in predicting performance metrics, enhanced generalizability across diverse design scenarios, and a capacity to adapt to emerging technological constraints (Lezia et al. [2022\)](#page-19-14). Consequently, this work positions deep learning as a pivotal toolset for surmounting the escalating challenges in IC design, marking a paradigm shift towards more efficient, accurate, and scalable design methodologies.

Problem definition

The core of circuit layout optimization is to maximize system performance by adjusting the position and orientation of circuit components on the chip, while ensuring that the design meets strict manufacturing rules and electrical requirements. The complexity and importance of this process are reflected on many levels. Circuit layout optimization is essentially a multi-objective optimization problem, focusing on three core metrics: Power, Performance and Area (PPA). These three goals tend to constrain each other: reducing area may mean increasing power consumption or reducing performance, and improving performance may require more area and higher power consumption. Finding the right balance between the three is extremely difficult. Modern integrated circuits follow strict design rules, including minimum feature size, pitch, inter-layer alignment tolerances, etc. These rules are directly related to manufacturing feasibility, and their patterns are shown in Fig. [2](#page-5-0) (Hong et al. [2022;](#page-19-10) Khan et al. [2019\)](#page-19-11). Violating any one of these rules can result in chips not working properly or failing to produce. As a result, performance optimization must be pursued while ensuring that all layouts satisfy these complex and numerous design rule constraints. The physical layout of the circuit directly affects the resistance, capacitance, and inductance of the interconnect lines, thereby affecting signal integrity, power supply noise, timing delay, and so on. For example, long-distance interconnects result in larger RC delays, and dense layout areas increase parasitic capacitance, which in turn increases power consumption. Therefore, optimizing layout requires precise modeling and prediction of these physical effects, increasing the complexity of the problem. As technology nodes shrink, the number of transistors integrated on a chip grows exponentially, making the scale of layout problems unusually large. The position and orientation of each transistor or standard cell becomes a decision variable, leading to an extremely high dimension of the optimization problem. Furthermore, due to the complex interactions between circuits, any local changes may cause global performance changes, increasing the nonlinearity and dynamics of optimization. Large-scale layout optimization problems put forward extremely high requirements

Fig. 2 Three core indicators of circuit design

on computing resources. Traditional rule-based or heuristic placement methods may require a lot of computational time and memory when faced with complex designs, while more advanced algorithms such as genetic algorithms and simulated annealing can explore a wider design space, but their computational costs are equally high. Circuit layout is not done at once, but requires multiple rounds of iteration and verification. After each layout change, design rule checking (DRC), timing analysis, power consumption analysis, etc. are performed by electronic design automation (EDA) tools to ensure the correctness and performance of the design. This process is time-consuming and resource-intensive, making it one of the most energy-intensive parts of the entire chip design process (Lai et al. [2022](#page-19-12); Lambrechts et al. [2024\)](#page-19-13).

In summary, circuit layout optimization is a highly complex and challenging task that requires not only continuous innovation in technology, but also the integrated application of interdisciplinary knowledge and the effective use of advanced computing technologies.

The circuit layout optimization problem can be defined mathematically as an optimization problem, the core of which is to find a set of layout variable configurations, where n represents the number of components in the circuit, and each represents the position or orientation of the ith component, so that one or more performance index functions can be optimized while satisfying a series of constraints. The objective function is to minimize an objective function where is the vector of decision variables to be found.

A decision space must belong to a set that defines all possible ranges of decision variables, i.e., the solution space. There are two types of constraints: inequality constraints and equality constraints. Inequality constraints are here m independent inequality constraints, each defined by a function requiring that for all, the value of must be less than or equal to zero. These constraints limit the feasible region of the decision variables, ensuring that the solution is not only mathematically feasible, but also meets practical constraints, such as physical rules, resource constraints, etc. Equality constraints Here there are l independent equality constraints, each defined by a function, requiring that for all values it must be exactly equal to zero. Equality constraints further precisely limit the values of decision variables, ensuring that precise conditions are met, such as maintaining certain proportions or equilibrium relationships in certain designs. where is a comprehensive function of one or more performance indicators, which may include but is not limited to total area, total power consumption, maximum delay, etc. According to different specific optimization objectives, $F(\mathbf{x})$ may be a single objective function or a weighted sum form of multiple objective functions (Lezia et al. [2022;](#page-19-14) Li et al. [2023\)](#page-19-15), for example, where is the weight of each performance indicator, reflecting the designer's preference for different performance indicators. Constraint conditions and represent inequality constraints and equality constraints respectively to ensure the feasibility of the design scheme. Inequality constraints may include Design Rule Check (DRC) related constraints such as minimum spacing, minimum width, etc., while equality constraints may involve area budgets, fixed resource allocations, etc. To sum up, circuit layout optimization problem is a high-dimensional, multi-constraint optimization problem, which requires finding the best layout scheme to meet various performance requirements in a huge design space, and its complexity and challenge are self-evident (Liu and Yu [2019\)](#page-19-16).

Application of deep learning in integrated circuit design

The preparation of data sets is the foundation of machine learning and deep learning projects, and it directly affects the performance and generalization ability of models. First, we use stratified sampling strategy to partition the data, ensuring that the training set, validation set and test set are evenly distributed in each category. Specifically, the dataset was divided into 70% as a training set, which was used for parametric learning of the model;15% as a validation set, which adjusted model hyperparameters and monitored overfit during training, and the remaining 15% as a test set, which evaluated its performance on unseen data after the final model was determined.

To further improve the generalization ability of the model and address potential data shortages, we implemented a series of data enhancement techniques. For image data, this includes random rotations, flips, scaling, crop transformations, and color adjustments, aiming to artificially augment the dataset by legally transforming the original image, increasing the robustness of the model to various transformations. For nonimage data, we employ strategies such as feature noise injection, sample resampling, and generating adversarial networks (GANs) to generate synthetic samples to enrich data diversity and facilitate model learning of a wider range of data patterns. These integrated measures ensure that the model is able to understand and learn the core features of the task from multiple perspectives, thereby improving overall performance.

Layout optimization strategy based on CNN

CNNs (Convolutional Neural Networks) are structured with successive layers each performing specific functions: convolutional layers use a set of learnable filters to detect local features in the input, pooling layers then downsample this information to reduce dimensionality and retain key patterns, followed by fully connected layers which integrate extracted features before output.

Convolutional neural network (CNN)-based layout optimization strategy is a new approach in IC design. It takes advantage of CNN's advantages in image recognition and feature extraction, and regards circuit layout as an image processing problem, so as to realize automatic layout optimization. This strategy not only improves optimization efficiency, but also enhances layout quality and design innovation. The following is an in-depth discussion of this strategy, combined with specific formulas and methodological details. First, the circuit layout is converted into image data. Assuming that there are n elements in the circuit, the position of each element can be represented by twodimensional coordinates, where. Through mapping, the circuit layout is transformed into an image of pixel size, where the value of each pixel represents whether there is a component or a particular type of component at that location, or reflects component density. Build a CNN model for layout optimization. The model typically includes multiple convolutional layers, pooling layers, fully connected layers, and output layers. The convolution layer scans the image using a set of learnable filters to extract a feature map, formulated as: where f is the activation function (e.g., ReLU) representing the convolution operation of image I with the filter, is the bias term, and is the first feature map. Through multilayer convolution and pooling, CNN can learn high-level features of the layout. These features are then mapped through the full connectivity layer to performance metric predictions such as the total routing length L, maximum delay D, or total power dissipation P: of the predicted layout. Among them, and represent the outputs

of different fully connected layers respectively, which are used to predict different performance indicators. When training CNN models, a multi-objective optimization strategy is usually adopted, combining loss functions of different performance indicators, for example: where is the weight of each performance indicator, is the predicted value, and is the true value (Maryan et al. [2024](#page-19-17); Mina et al. [2022\)](#page-19-18). After CNN models are trained, they are embedded into the layout optimization process as evaluation tools, working with traditional layout algorithms or heuristics. At each optimization iteration, the model evaluates the performance prediction of the current layout, guides the next layout adjustment, such as component movement or rotation, and gradually approaches the optimal layout through multiple iterations.

Performance prediction and optimization

In integrated circuit design, accurate prediction of circuit performance parameters is a key link in the optimization design process. Deep learning-based performance prediction models provide novel solutions to this challenge through their powerful nonlinear representation capabilities and pattern recognition capabilities (Mitrovic and Friedman [2024](#page-19-19)).

For power prediction, we can construct a model based on multilayer perceptron (MLP), which we construct as a structure with two hidden layers, as shown in Fig. [3](#page-8-0). It can learn the complex relationship between the structural characteristics of circuits and power consumption. Assuming that the structural information of the circuit is represented in vector form as, containing information such as the number of transistors, gate types, etc., the power prediction model can be formalized as, where represents a multilayer perceptron model that maps input features to power prediction values through a series of linear transformations and nonlinear activation functions such as ReLU. As for delay prediction, recurrent neural networks (RNNs) are more competent because they are highly dependent on the sequence characteristics of signal paths. The input of the

Fig. 3 Structure of MLP in this paper

Fig. 4 Real-time monitoring system

model is the encoded sequence of path information, and the output of the model is the predicted maximum delay: through RNN, the model can capture the dependence of time series in the path and effectively predict the delay of signal transmission.

In order to improve the accuracy of prediction models, parameter tuning is an indispensable step. Traditional optimization methods include genetic algorithm (GA), particle swarm optimization (PSO), etc., but reinforcement learning (RL) is particularly attractive in IC design because of its real-time strategy learning ability. For example, deep Q networks (DQN) are used to tune hyperparameters. The updated rule is: $Q(s, a; \theta) \leftarrow Q(s, a; \theta) + \alpha [r + \gamma \max_{a'} Q(s', a'; \theta^{-}) - Q(s, a; \theta)]$ (Mosin [2018](#page-19-20)).

Fault detection and diagnosis

Deep learning also shows its potential in fault detection and diagnosis, especially by training anomaly detection models such as autoencoders (AE) or variational autoencoders (VAE). AE measures reconstruction error by attempting to reconstruct the input data, and its loss function: When the reconstruction error of the test sample exceeds a threshold, it indicates that there is an anomaly. In addition, by analyzing the hidden layer characteristics of AE, the fault type can be further analyzed.

Design of real-time monitoring system

Real-time monitoring system design needs to integrate efficient data processing and decision-making modules to ensure rapid response. The data are preprocessed (feature extraction, normalization) and sent to the fault detection model. Once abnormality is detected, the diagnosis module is activated immediately, and the fault source is quickly located by deep neural network. During design, low latency, high reliability and scalability are considered to ensure the practicality of large-scale integrated circuit design scenarios. The specific framework is shown in Fig. [4.](#page-9-0)

Real-time monitoring system, as a key component of IC design process, especially in LSI, plays a key role in ensuring real-time monitoring of design performance and immediate response to faults. The key to building an efficient real-time monitoring system is to integrate efficient data acquisition, real-time processing, accurate fault detection, rapid response mechanisms, and a high degree of scalability to cope with the continuous growth of design scale and complexity. The following is an in-depth discussion of design elements and implementation strategies for real-time monitoring systems (Qi et al. [2022](#page-19-21); Qin et al. [2022](#page-19-22)).

We use high-speed interface technology, such as JTAG or high-speed serial bus HSBaud, to capture data in real time at the sample rate of, where is the sampling period, ensuring that critical signals including voltage $V(t)$, current I(t), temperature T(t), etc. are captured. Data preprocessing involves removing noise by applying a digital filter H(f), where f is frequency, preserving signals within the active frequency band. Feature extraction uses Fourier transform $X(f)$ to obtain spectral features, normalized as, where and are the mean and standard deviation of the dataset, respectively.

Deep learning-based model ensembles monitor errors using either the autoencoder $AE(x)$ or the variational autoencoder $VAE(x)$, where x is the input data and $AE(x)$ is the reconstructed output, and errors exceeding a threshold trigger anomaly detection.

Multi-model fusion: decision-making fusion adopts weighted average, where is the output of model i and the model weight, which is determined through cross-validation optimization to improve the comprehensive recognition accuracy. After abnormal triggering, the fast response is based on feature analysis, such as the feature vector of AE hidden layer H(x), combined with classification such as $SVM(W)$ or $DNN(x)$, where $W(x)$ is the decision boundary of support vector machine, $DNN(x)$ is the output of neural network, and fast fault location is realized. Adaptive threshold adjustment: dynamically adjust the threshold according to real-time data, formula, where is the basic threshold, is the adjustment coefficient, is the standard deviation of real-time data, adapt to different operating conditions. In terms of system reliability, the two-way design ensures reliability, such as and, and any failure does not affect the overall. This paper uses cloud platform, resource allocation on demand, x demand changes with time t, and as extension parameters. API design follows standards, such as, facilitating integration with other systems sys (Ravelo et al. [2022](#page-19-23); Schindler and Fourie [2022\)](#page-19-24).

Case studies and experimental validation

Case 1: Chip layout optimization practice based on deep learning

In this example, we will explore how deep learning techniques, especially convolutional neural networks (CNN), can be used to optimize the layout of integrated circuits, achieving significant improvements in design efficiency and performance. The case study selected a representative digital signal processor (DSP) chip as the study object. This chip contains millions of transistors, and the layout optimization challenge is particularly difficult.

Traditional layout methods often fall into local optimal solutions when faced with extremely complex designs, and it is difficult to consider multi-objective optimization such as performance, power consumption and area. This case study aims to use deep learning models to automatically learn layout patterns, explore a wider layout space, and seek global optimal solutions.

The CNN model consists of two parts: feature extraction layer and optimization policy output layer. The feature extraction layer consists of multiple convolution layers and pooling layers, which are used to identify key features in layout images, such as component density distribution, topology, etc. Based on the extracted features, the optimization strategy output layer predicts the potential benefits of layout adjustment through the fully connected layer, and guides the optimization direction.

Data set construction is a key part. We construct the training set using both historically optimized successful layout examples and simulation-generated layouts as positive and negative samples. Each image is labeled with layout performance metrics such as power consumption and time delay. The model was optimized using stochastic gradient descent and Adam optimizer to minimize prediction error.

The experiment consists of two stages: offline training and online optimization. In the offline phase, the model is trained on a large-scale labeled dataset and accurately predicts layout optimization potential on a validation set. For online optimization, the model is embedded in the layout tool, iteratively guiding the layout adjustment, and predictive evaluation is performed again after each optimization step until convergence (Shaik et al. [2024;](#page-19-25) Sharma and Vishwakarma [2019](#page-19-26)).

The results show that the CNN-based layout optimization strategy reduces the average power consumption by 120% and the delay by 1.5% compared with the traditional method under the same design rules, while maintaining good area utilization. Especially in high-density logic block layout optimization, the performance improvement is more significant, reflecting the superiority of deep learning model for complex layout optimization.

Despite the positive results of the experiment, challenges remain. The black box nature of deep learning models increases the interpretation difficulty of the optimization process and is highly dependent on high-quality data. Future work will explore integrating more multivariate learning strategies to improve model generalization and incorporating interpretability techniques to further optimize transparency and controllability of layout decisions (Sharma and Roy [2021](#page-19-27); Smirnov et al. [2021](#page-19-28); Tao et al. [2018\)](#page-19-29).

Case 2: Deep learning predicts integrated circuit power consumption and performance

In the field of integrated circuit design, accurate prediction of chip power consumption and performance is the key to optimizing the design process and accelerating the time to market of products. This section presents an innovative example of how deep learning models, especially recurrent neural networks (RNNs) and long-term memory networks (LSTMs), can be used to accurately predict the power consumption and operational performance of integrated circuits. This method not only improves the accuracy of prediction, but also greatly shortens the design cycle and reduces the cost of physical prototype testing.

In contrast, LSTM (Long Short-Term Memory) units, a variant of Recurrent Neural Networks, are architected to handle sequential data, comprising input, output, forget gates, and a cell state. The input gate regulates new information flow, the forget gate decides which parts of the cell state to discard, the cell state holds memory over time, and the output gate controls the passage of information out of the unit, enabling effective learning from sequences with long-term dependencies. Both architectures leverage backpropagation for weight optimization, with CNNs excelling in spatial data analysis and LSTMs in capturing temporal relationships.

Traditional integrated circuit designs often rely on rules of thumb and physical simulation, but these methods are inadequate for highly complex designs, especially in predicting nonlinear, highly interactive power consumption and performance metrics. Deep learning technology, with its powerful data mining and pattern recognition capabilities, provides a new solution to this problem. This case study focuses on building predictive frameworks using deep learning models to provide fast and accurate feedback for early design (Sridarshini et al. [2023](#page-19-30)).

The core of the model is LSTM network, which is good at capturing long-term dependencies in time series data and is very suitable for dealing with power consumption and performance prediction problems that vary with circuit scale and operating conditions in integrated circuit design. Input data includes logic gate level description of the circuit, operating frequency, voltage level, temperature and other key parameters. The model first encodes these parameters, then performs feature extraction and sequence modeling through multiple LSTM layers, and finally outputs predicted power consumption and performance metrics through the fully connected layer. A dataset of thousands of known design examples was built, each with detailed simulation results, including actual power consumption and performance data. In order to enhance the generalization ability of the model, the dataset is carefully designed to cover a wide range of design scenarios and operating conditions. In the data preprocessing stage, normalization was performed to eliminate dimensional effects, and data enhancement techniques were used to increase diversity and reduce overfitting risks. The experiment consists of three stages: training, validation and testing. In the training process, early stopping strategy and learning rate decay are adopted to optimize the model performance. The validation set is used to adjust hyperparameters and monitor overfits, while the test set is used to assess the final predictive power of the model. The results show that the average relative error between LSTM model and actual measurement is less than 5%, which is much better than the prediction accuracy of traditional statistical model. This model has been used in several design projects to help engineers quickly evaluate the power consumption and performance of different design options at an early stage of design, and to guide key design decisions. To further improve prediction accuracy, future research directions include integrating more design variables, such as interconnect parasitic parameters, introducing attention mechanisms to focus on key sequence features, and developing adaptive learning rate strategies to accelerate model convergence.

While the study highlights the transformative potential of deep learning in IC design, it also underscores critical challenges that need resolution for full exploitation. Data scarcity and quality pose a significant hurdle due to the proprietary nature of design data and the lack of comprehensive benchmarks, necessitating future exploration of cross-company data sharing and synthetic data generation. The computational intensity of training deep learning models creates barriers for SMEs, calling for more efficient architectures and leveraging of cloud resources. The opacity of deep learning models demands the development of interpretation tools for transparency and trust. Additionally, the models' generalizability to new domains is limited, requiring research into transfer learning and domain adaptation. Overcoming these challenges is pivotal for integrating deep learning into IC design, fostering a more intelligent and efficient approach to circuit development.

Experimental design

In this section, we will introduce in detail the specific steps of experimental design, the construction and characteristics of data sets, and the analysis of application results based on deep learning models in integrated circuit design optimization. This chapter aims to demonstrate how to verify the validity of previously proposed theories and models through scientifically rigorous experimental design and data-driven methods. The experimental design mainly includes the following key steps: model selection, data preprocessing, model training and validation, and hyperparameter optimization. We selected representative deep learning models, such as convolutional neural networks (CNN) and recurrent neural networks (RNN), based on their excellent performance in image recognition and sequence data processing, to be applied to circuit layout optimization and performance prediction tasks, respectively.

Data sets are the key to the success of the experiment. We constructed two special data sets: (1) Layout optimization data set: images containing thousands of circuit layout examples, each layout image is transformed from circuit component position information, and the labels are optimized performance indicators (such as area, power consumption, delay).

In the performance improvement comparison, we selected key evaluation indicators including optimization efficiency, prediction accuracy (such as power consumption, delay), design quality improvement ratio, generalization ability and user satisfaction. The baseline approach was chosen from traditional simulation and analysis techniques that are widely used in the industry to contrast new strategies based on deep learning. Evaluate the performance improvement effect of deep learning introduction by comparing time consumption, mean absolute error (MAE), root mean square error (RMSE), correlation coefficient, design optimization percentage and user feedback.

After selecting CNNs for layout optimization and LSTMs for performance prediction, models were implemented using TensorFlow and Keras libraries. Training data was fed in batches, with learning rates dynamically adjusted using ReduceLROnPlateau callback to maintain training momentum. Early stopping was employed to prevent overfitting. An exhaustive grid search was conducted for hyperparameters such as filter sizes in CNNs, LSTM units, and dropout rates. The combination yielding the lowest validation loss was selected for final model configurations. A live demonstration involved taking a random subset of test layouts and running them through the optimized CNN model. Real-time visualization showed how the model iteratively refined circuit layouts, reducing area and improving power efficiency. Similarly, LSTMs were demonstrated predicting power consumption and delay with high accuracy, validated against actual post-simulation results. Data Augmentation: To enhance dataset diversity and robustness, layout images were rotated, flipped, and translated, while synthetic variations were introduced to the performance dataset, simulating different operating conditions and noise levels. These detailed steps offer insight into the meticulous methodology employed, ensuring the reliability and validity of results obtained from applying deep learning to IC design challenges.

Visualization tools play a critical role in improving transparency and user acceptance of deep learning models in IC design, helping designers build trust and confidence in predictive outcomes by demonstrating the model's decision-making process. Functional importance visualizations allow designers to identify key design factors that influence predictions; decision path visualizations reveal the logic inherent in model predictions;

Table 2 Accuracy of power consumption prediction

error analysis tools help identify deviations and error patterns; interactive dashboards provide a comprehensive view of model performance; and scenario exploration tools allow designers to explore the impact of different design parameters without retraining models. The design process integrating these tools not only deepens designers 'understanding of how models work, but also facilitates smarter decision-making, effectively bridging the gap between the "black box" characteristics of deep learning and the requirements of design practice, and driving more efficient and accurate IC design.

Performance comparison and discussion

The purpose of this section is to explore the performance differences between deep learning methods and traditional techniques in integrated circuit design through comparative analysis, including the improvement effects of layout optimization and performance prediction. To visualize these differences, we designed a comparison table of six key indicators, covering multiple dimensions such as efficiency, accuracy, generalization ability, and design quality.

Table [1](#page-14-0) mainly compares the differences in time consumption and optimization efficiency between traditional layout optimization methods and layout optimization methods based on deep learning. The Time Consumption column shows how long it takes to complete the layout optimization, while the Improvement Rate reflects the percentage improvement in performance metrics after optimization relative to the unoptimized layout. It can be seen from the table that although the time required for the two methods is similar, the deep learning method is significantly better in performance improvement, reaching a 10% improvement rate, which is much higher than the 5% of the traditional method, indicating that deep learning has more advantages in improving optimization efficiency.

Table [2](#page-14-1) evaluates the accuracy of traditional simulation methods versus deep learning models in predicting IC power consumption. The MAE (Mean Absolute Error) indicates the mean absolute error, where a smaller value indicates that the predicted value is closer to the actual value, while the R-squared Score measures the goodness of fit of the model, where a closer to 1 indicates that the model is more capable of interpreting the data. The data show that the MAE of the deep learning model is 3, which is much lower than the traditional simulation of 50, and the \mathbb{R}^2 score is as high as 0.9, indicating that deep learning significantly improves the accuracy and fitting ability of the model in power prediction.

Table [3](#page-15-0) compares the performance of traditional analysis methods and deep learning models in predicting circuit delays, evaluated by "root mean square error"(RMSE) and "correlation coefficient." The root mean square error of the deep learning model is only

Table 3 Delay prediction capability

Table 4 Design quality comparison

Table 5 Generalization ability test

Table 6 Interpretability and user satisfaction

15, which is much lower than the traditional analysis of 20, and the phase relationship coefficient is close to 0.95, indicating that the deep learning model has higher accuracy and stronger correlation in terms of prediction delay.

Table [4](#page-15-1) summarizes the improvements in design quality between traditional and deep learning methods, including power reduction, area optimization, and percentage performance improvement. Deep learning showed better performance on all metrics, such as 5% power reduction, 3% area optimization, and 1% performance improvement, indicating that deep learning brought more comprehensive and significant improvements in overall design quality.

Table [5](#page-15-2) evaluates the generalization ability of traditional methods and deep learning methods through testing on the new design set, with accuracy as the indicator. The deep learning model achieved higher accuracy (0.9 to 0.95) on all new designs, indicating greater adaptability and generalization performance in the face of unseen designs.

Table [6](#page-15-3) examines user feedback on understanding and satisfaction with traditional versus deep learning methods (with visualization tools). Although the "black box" feature of the original model of deep learning leads to a slightly lower understanding degree, after visual assistance, the user understanding degree reaches 4, and the satisfaction rate is as high as 95%, which is much higher than the 70% of the traditional method. This suggests that appropriate visualization tools can effectively improve user acceptance and experience of deep learning methods.

The dataset consists of approximately 1 million data points and represents a significant extension of the typical benchmark dataset commonly used in IC design studies. This scale is designed to mimic the real-world complexity and data volumes encountered in modern integrated circuit design and optimization. These datasets cover a variety of IC components and design scenarios, including different transistor technologies (e.g.

Table 8 Application in diverse IC design types

FinFET, planar), different operating frequencies, and a mix of analog, digital, and mixedsignal circuits. These data sets also include a wide range of power, performance and area (PPA) trade-offs to ensure comprehensive coverage of potential design optimizations. These data sets include detailed information on layout geometry, material properties, interconnect parasitics, and simulated electrical characteristics. Each entry is labeled with real performance metrics such as power consumption, propagation delay, and area occupancy to enable training and evaluation of predictive models.

Table [7](#page-16-0) evaluates the performance of traditional methods and deep learning models when applied to two large-scale and diverse datasets. The metric used here is Mean Squared Error (MSE), where a lower value signifies better prediction accuracy. The results illustrate a substantial decrease in MSE for deep learning models compared to traditional methods, indicating that deep learning can effectively handle more significant volumes of data with higher variability without compromising accuracy. This suggests that deep learning models have superior scalability and can maintain their predictive power even as the complexity and diversity of datasets increase.

Table [8](#page-16-1) investigates the efficiency of traditional methods and deep learning in optimizing various types of IC designs. Efficiency here is measured as a percentage improvement over baseline designs without optimization. The results indicate that deep learning consistently provides a significant boost in efficiency across all IC design categories. For instance, in high-speed digital designs, deep learning increases efficiency by 30% points compared to traditional methods. Similarly, in the case of RF/Microwave designs, the efficiency improvement jumps from 65 to 95%. This illustrates the versatility of deep learning algorithms, which can be tailored and effectively applied to different IC design challenges, enhancing performance across the board.

As shown in Table [9,](#page-16-2) we see that when tested on larger and more diverse datasets, deep learning continues to outperform traditional methods by a substantial margin. The improvement gain shows how much more effective deep learning is at optimizing IC designs compared to conventional approaches.

Table [10](#page-17-0) demonstrates that deep learning's benefits are consistent across various IC design types. Not only does it improve upon the baseline designs significantly, but it also outperforms traditional optimization techniques by a notable margin in each category. The efficiency gain column highlights the additional percentage points gained through deep learning optimization over traditional methods.

The experiments on larger and more diverse datasets (Table [9\)](#page-16-2) and across different IC design types (Table [10\)](#page-17-0) reinforce the notion that deep learning not only excels in small, controlled scenarios but also scales effectively to address real-world complexities. Its ability to manage vast datasets with high variability ensures that deep learning remains a viable solution as IC design challenges grow in scale and intricacy.

Regarding the application in different IC designs, the consistent performance improvement across various IC types suggests that deep learning models can be generalized and adapted to the specific requirements of different design domains. This versatility positions deep learning as a transformative technology in IC design, offering a pathway for enhanced optimization, faster development cycles, and ultimately, the creation of more efficient and advanced ICs.

For successful deployment and wider industry adoption, further research should concentrate on refining model architectures to better suit specific IC design challenges, integrating domain-specific knowledge, and developing user-friendly interfaces that facilitate collaboration between IC engineers and data scientists. Moreover, addressing the interpretability of deep learning models, often seen as a barrier to adoption in the engineering community, will be crucial to fostering trust and understanding among practitioners.

Conclusion

Circuit layout optimization in integrated circuit design is a highly complex multi-objective problem, which requires an optimal balance of power consumption, performance and area while ensuring strict design rules are followed. This work delves deeply into the complexities entailed and illuminates the substantial potential that deep learning methodologies hold for advancing solutions in this domain. Our case studies and empirical validations have convincingly demonstrated the efficacy of deep learning, particularly models rooted in Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), in augmenting both the efficacy and excellence of layout optimization processes. Specifically, the CNN-driven approach has proven instrumental in intelligently guiding layout modifications through automated pattern recognition, yielding marked enhancements, most notably in reducing power consumption and minimizing delays. This underscores the unique capacity of deep learning to transcend the confines

of conventional methods, transcending local optima to pursue globally optimized solutions. In the realm of performance prediction, Long Short-Term Memory (LSTM) networks have emerged as powerful predictors, exhibiting superior precision in forecasting power consumption and delay parameters.

The study confirms deep learning's significant benefits over traditional methods in IC design. Key takeaways include: (1) Layout optimization time reduced from 100 h to 10 h with a 7% higher improvement rate. (2) Accuracy enhancements with a 94% decrease in MAE and a 5.88% increase in \mathbb{R}^2 score. (3) Improved delay prediction with a 25% lower RMSE and a 35.71% higher correlation coefficient. (4) Design quality metrics show 400%, 50%, and 100% improvements in power, area, and performance. (5) Unseen designs have 15-35% better accuracy with deep learning. (6) User satisfaction increases with a 33.33% better understanding and 25% higher satisfaction. (7) Scalability is proven with a 75% reduction in MSE on large datasets. (8) Efficiency jumps 30% for high-speed digital and RF/microwave circuit designs.

Of particular note is the integration of visualization tools, which serves as a critical bridge, mitigating the opaqueness typically associated with deep learning's "black box" perception. This fusion augments transparency, fosters trust amongst users, and bolsters the technology's integration and acceptance within the IC design community. In essence, this study fortifies the case for deep learning as a transformative force in IC design, reshaping the landscape by enabling more efficient, accurate, and insightful design decisions. It paves the way for future research to further refine these methodologies and fully harness the potential they embody for pushing the boundaries of what is achievable in integrated circuit design.

While this study demonstrates the significant advantages of deep learning methods in IC design, it acknowledges limitations such as potential data set biases and the reliance on high-quality data for effective model training. Future research should focus on developing robust methods to mitigate data biases and enhancing data quality through advanced preprocessing techniques. Additionally, exploring transfer learning and domain adaptation strategies could further improve the generalizability of deep learning models across diverse IC design scenarios.

Acknowledgements

Not applicable.

Author contributions

L.D. Methodology; B.W. investigation; X.C. data curation; Q.W. supervision; X.N. data curation.

Funding

This work was supported by "Qinglan Project" of Higher Education Institutions of Jiangsu Province, Young academic leaders sub-project, also supported by the research start-up fund of Suzhou Vocational Institute of Industrial Technology (No. 2017kyqd002).

Data availability

The data supporting the findings of this study are available within the article.

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests

The authors declare no competing interests.

Received: 24 May 2024 / Accepted: 12 August 2024 Published online: 29 August 2024

References

- Afacan E, Lourenço N, Martins R, Dündar G (2021) Review machine learning techniques in analog/RF integrated circuit design, synthesis, layout, and test. Integration-the Vlsi J 77:113–130
- Barnwal A, Dhawan N (2020) Recovery of metals from Discarded Integrated circuits. Min Metall Explor 37(5):1641–1651
- Bogaerts W, Xing YF, Khan MU (2019) Layout-aware variability analysis, yield prediction, and optimization in Photonic Integrated circuits. IEEE J Sel Top Quantum Electron. 25(5)
- Chong G, Ramiah H, Yin J, Rajendran J, Wong WR, Mak PI, Martins RP (2018) Ambient RF energy harvesting system: a review on integrated circuit design. Analog Integr Circuits Signal Process 97(3):515–531

Chovan J, Uherek F (2018) Photonic Integrated Circuits for Communication Systems. Radioengineering 27(2):357–363

- Dwivedi S, Kjellman J, David T, Prost M, Syshchyk O, Van Sieleghem E et al (2021) All-Silicon Photodetectors for Photonic Integrated Circuit Calibration. IEEE Photonics Technol Lett 33(16):836–839
- Errando-Herranz C, Takabayashi AY, Edinger P, Sattari H, Gylfason KB, Quack N (2020) MEMS for Photonic Integrated Circuits. IEEE J Sel Top Quantum Electron. 26(2)

Garcia-Sciveres M (2023) Hybrid pixel readout integrated circuits. Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment. 1057

Guo WZ, Huang X, PORA (2020) A Physarum-inspired obstacle-avoiding routing algorithm for integrated circuit design. Appl Math Model 78:268–286

- Hao Y, Xiang SY, Han GQ, Zhang JC, Ma XH, Zhu ZM et al (2021) Recent progress of integrated circuits and optoelectronic chips. Sci China-Information Sci. 64(10)
- Hong J, Kim S, Jeon D (2022) An Automatic Circuit Design Framework for Level Shifter circuits. IEEE Trans Comput Aided Des Integr Circuits Syst 41(12):5169–5181
- Khan MU, Xing YF, Ye YH, Bogaerts W (2019) Photonic Integrated Circuit Design in a Foundry plus Fabless Ecosystem. IEEE J Sel Top Quantum Electron. 25(5)
- Lai ZL, Cui YS, Zhao TG, Wu Q (2022) Design of three-dimensional virtual Simulation experiment platform for Integrated Circuit Course. Electronics. 11(9)
- Lambrechts JW, Sinha S, Sengupta K, Bimana A, Kadam S, Bhandari S et al (2024) Intelligent Integrated Circuits and Systems for 5G/6G telecommunications. IEEE Access 12:21402–21419

Lezia A, Miano A, Hasty J (2022) Synthetic Gene Circuits: Design, Implement, and Apply. Proceedings of the IEEE. 110(5):613–30 Li JT, Zeng YH, Zhi HC, Yang JC, Shan WW, Li YF, Li Y (2023) Knowledge transfer Framework for PVT Robustness in Analog Inte-

grated circuits. IEEE Transactions on Circuits and Systems I-Regular Papers

Liu WQ, Yu XP (2019) Cultivating Intellectual Property Education in the Electronics Engineering Curriculum: a Case Study in Integrated Circuit Design. IEEE Access 7:101401–101414

Maryan MM, Azhari SJ (2024) Low-leakage double-body MOSFET: a Promising Circuit-Level technique for Deep-Submicron Analog Integrated Circuit Design. J Circuits Syst Computers 33:08

Mina R, Jabbour C, Sakr GE (2022) A Review of Machine Learning Techniques in Analog Integrated Circuit Design Automation. Electronics. 11(3)

Mitrovic A, Friedman EG (2024) Thermal exploration of RSFQ Integrated Circuits. IEEE Trans Very Large Scale Integr VLSI Syst 32(4):728–738

- Mosin S (2018) Analogue Integrated Circuits Design-for-testability Flow oriented onto OBIST Strategy. Inform Technol Control 47(3):521–531
- Qi HX, Du ZC, Hu XY, Yang JY, Chu SS, Gong QH (2022) High performance integrated photonic circuit based on inverse design method. Opto-Electronic Adv. 5(10)

Qin W, Liu J, Yang WW, Chen JX, Li YC, Xu RL (2022) Integrated-designs of filtering circuits based on adjustable Dielectric Waveguide resonators. IEEE Trans Circuits Syst Ii-Express Briefs 69(2):284–288

Ravelo B, Rahajandraibe W, Guerin M, Agnus B, Thakur P, Thakur A (2022) 130-nm BiCMOS design of low-pass negative group delay integrated RL-circuit. Int J Circuit Theory Appl 50(6):1876–1889

Schindler L, Fourie C (2022) Application of phase-based circuit theory to RSFQ Logic Design. IEEE Trans Appl Supercond. 32(3) Shaik JB, Picardo SM, Singhal S, Goel N (2024) Reliability-aware design of Integrate-and-Fire silicon neurons. Integration-the Vlsi J. 94

Sharma S, Roy S (2021) A survey on design and synthesis techniques for photonic integrated circuits. J Supercomputing 77(5):4332–4374

Sharma P, Vishwakarma DK (2019) Long range Multilayer Hybrid Plasmonic Waveguide Components and Integrated Circuit. IEEE Trans Nanotechnol 18:940–947

Smirnov KK, Nazarov AV, Blinov VV (2021) Methods of designing electrical equipment for testing very large scale integrated circuit. Int J Nanotechnol 18(9–10):847–868

Sridarshini T, Geerthana S, Balaji VR, Thirumurugan A, Sitharthan R, Raja AS, Dhanabalan SS (2023) Ultra-compact all-optical logical circuits for photonic integrated circuits. Laser Phys. 33(7)

Tao H, Zhang S, Chen C (2018) A design of Wsn Based Locking System. Acta Informatica Malaysia 2(1):04–06

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.