

人工智能会改变 EDA 吗？

0. 前言

近年来，人工智能（AI）与机器学习（ML）已经在诸多应用领域取得了突破。而在传统的半导体领域，研究者们也探索了基于机器学习的芯片设计新方法。这些新算法通常会最先反映在芯片设计工具上，也就是我们常说的 EDA 工具。因此，这一研究方向通常可以被称为 ML for hardware design，或者 ML for EDA（严格来说前者的范围更大一些）。有时也称做智能化的 EDA 算法或者智能的 IC 设计方法。

本文将介绍 ML for EDA 这一 EDA 领域的热门研究方向。这一方向涉及到机器学习，数据结构与算法，电路设计与制造等基础知识，是典型的交叉研究方向。我们将从 EDA 的背景知识开始，介绍什么是 ML for EDA，为什么我们研究 ML for EDA，该研究方向早期一些代表性的工作，当前商业化的进程，其他相关的研究方向，以及 ML for EDA 可能面临的挑战。

1. 什么是 EDA? – EDA 与芯片设计基础

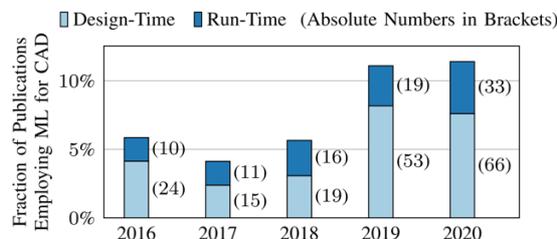
芯片在我们的生活中无处不在。芯片的设计与实现涉及一个复杂的流程。以数字芯片为例，假设设计团队从已经完成的 Verilog/VHDL 代码出发，标准设计流程至少还要包括逻辑综合（logic synthesis），布局规划（floorplan），时钟树综合（CTS），布局布线（placement & routing）等步骤。除此之外我们还需要进行大量的仿真和验证工作。在这个过程中，工程师需要权衡功耗，频率，面积等多个设计目标，同时还要确保制造出的芯片将正确运行各种功能。另外随着摩尔定律的进展，当下大规模的芯片已经可以包含超过 100 亿个晶体管。考虑到这种流程与设计的复杂性，几乎所有设计团队都需要商业 EDA 工具来辅助完成整个芯片设计。因此，如果我们把整个半导体产业比作一座金矿，EDA 工具则也许可以被看做挖矿的铲子。在美国频繁限制半导体出口的当下，本身市值并不大的 EDA 产业越来越为人们所重视。

EDA 的全称为 Electronic Design Automation，即电子设计自动化。经过几十年的发展，EDA 工具已经被用于芯片设计与制造流程的方方面面。EDA 工具的质量将会影响最终芯片的质量（功耗，频率，面积）以及设计效率（产品上市时间）。因此，EDA 算法研究者的目标通常至少包括两方面，一是芯片优化效果（功耗，频率，面积），二是 EDA 工具本身的效率（总耗时=工具单次运行时间*使用该工具的次数）。

这里我们强调了 EDA 工具的使用次数，因为芯片设计不是一次完成的。通常工程师需要多次迭代各个设计步骤直至达成设计目标。每一次迭代，工程师可以修改工具参数或者芯片设计，并重新运行 EDA 工具，以期更好的芯片效果。对于大规模的芯片设计，仅执行部分设计流程每次即可花费数天至数周时间。因此，迭代次数越多，花费在 EDA 工具上的时间越多，最终流片以及产品上市日期就会越晚。后文将会介绍，很多 ML for EDA 工作的原理就是减少迭代次数，尽快让芯片质量收敛到设计目标。

2. 什么是 ML for EDA? – 预测与优化

近年来，我们看到越来越多的 ML for EDA 方向的探索。如下图 [1] 所统计，发表在代表性 EDA 会议（DAC, ICCAD, ASP-DAC, ...）与期刊（TCAD）的相关工作数量逐年增加。

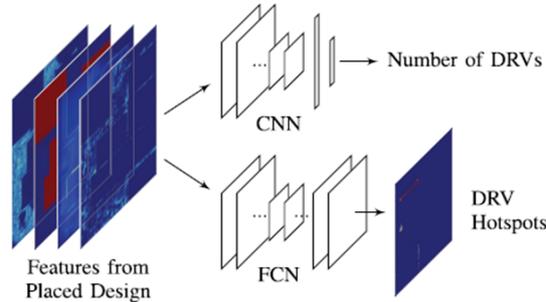


历年发表的 ML for EDA 论文数量呈增加趋势 [1]

这些 ML for EDA 的方法在芯片设计流程中的作用大体可以分为预测（prediction）与优化（optimization）两大方面。下面我们分别举例来介绍这两类应用。

预测类工作通常使用 ML 模型对芯片优化的最终目标进行早期快速预测。利用 ML 模型的预测，设计师可以减少对耗时的 EDA 工具的使用，而直接预测 EDA 工具运行后的大致效果。基于 ML 预测，设计者可以及时调整设计参数。换句话说，ML 模型通过快速预测 EDA 工具的行为，减少了运行 EDA 工具的次数。

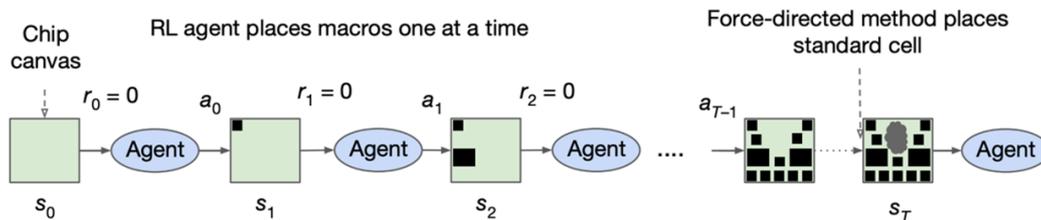
一个典型的预测类工作例子就是，使用 CNN 模型在芯片绕线之前对（绕线后将会产生的）DRC 热点（hotspot）进行提前预测 [2]。DRC 热点指的是绕线后违反设计规则的区域。通过预测，EDA 工具可以提前设法修改布局，避免未来阶段将会产生的 DRC 热点。这个工作可以类比于用 CNN 进行图像分类或者语义分割，其中芯片版图类似于图片，而需要预测的 DRC 热点位置类似于图片上需要分割的区域。



预测类方法举例：使用 CNN 对 DRC 热点进行预测 [2]

相比于预测类工作，优化类工作需要做得更多。这里说的优化是一种比较宽泛的概念，指的是直接解决一些 EDA 问题。事实上大量 EDA 问题本质都是对于芯片在一定约束下的优化问题，而很多传统 EDA 算法是大家多年积累的（对这些问题的）优秀启发式解法。那么优化类 ML 方法追求的是比传统 EDA 算法更优或者更快的解法。这个解法可以帮助生成真实精确的芯片设计结果。

一个典型的优化类例子就是谷歌使用强化学习（Reinforcement Learning）进行 macro 摆放（placement）[3]。通过强化学习算法，这个工作可以类比于 AlphaGo 进行围棋落子，其中芯片版图类似于围棋棋盘，而每一个 macro 元件的摆放位置类似于围棋每次最佳的落子位置。



优化类方法举例：使用强化学习进行 macro 摆放 [3]

按照这种不太严谨分类方式，预测类 ML 方法能够减少 EDA 算法的使用次数，从而提高芯片设计效率。而优化类 ML 方法能够帮助生成更快或更好的设计。因此如果想要彻底取代某些传统 EDA 算法，我们需要的是优化类方法。当然像布局布线这样包括百万元件的大规模优化问题，想直接取代传统 EDA 方法是困难的。目前我们看到更多的是 ML 方法融入 EDA 框架，起到重要的辅助性作用 [4]。

COMPARE ML AND EDA ALGORITHMS

	EDA Algorithms	ML
	Placement, route, synthesis, CTS, simulations, etc	Supervised, self-supervised, reinforcement learnings, etc
Pros	Known optimality, robust, less training data, good interpretability, Solve an abstracted problem efficiently	GPU parallel computing, easy to design, end-to-end training on complex problem, Solve any problem by learning from its data
Cons	Oversimplification of dynamic, complex problems	Rely too much on data, not leveraging the mechanics of the problem



Integrate ML with EDA Algorithms, Key is Problem decomposition

EDA 和 ML 方法的比较：应该将 ML 融入 EDA 方法 [4]

3. 为什么研究 ML for EDA ?

在深入介绍其他工作之前，一个需要回答的问题是，为什么我们要研究 ML for EDA ? 换句话说，ML for EDA 相比传统 EDA 算法有什么核心优势吗？这里尝试提供一种简单化的解释 [5]。

如背景中介绍，设计流程分为多个阶段。在大部分阶段，EDA 工具很难直接去优化最终的设计目标，因为最终的设计目标只有在整个流程走完之后才会确定。于是 EDA 工具只能对最终的设计目标进行一个粗略估计。举例来说，布局 (placement) 算法可能会优化总线长 (HPWL) 与拥塞 (congestion)，但这个优化目标这并不总能反映最终的设计目标 (最终芯片的功耗，频率，面积，DRC 热点，等等)。

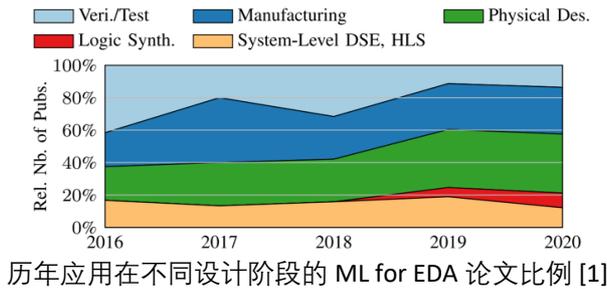
因此，除非每次都走完漫长的设计流程，不然前期的 EDA 工具并不能准确知道自己的解是否够好。为了保证设计收敛，一种策略是使用保守的早期估计，以给后期阶段留足余地 (margin)。但这显然会牺牲芯片质量。另一种策略是不断调整参数进行反复多次迭代，以期获得更好的芯片质量。而这非常依赖专业设计师的经验，并且会花费大量设计时间。

ML for EDA 算法的特点是它是由数据驱动的方法。通过学习已有的设计数据，早期使用 ML 预测可以获得更准确的优化目标，或者叫做早期反馈 (early feedback)。于是预测类 ML 方法打通了不同设计阶段之前的障碍，也减少了对人工的依赖。而对于优化类 ML 方法，强化学习类型的方法通过探索巨大的设计空间，有可能获得比传统启发式算法更优的解。另外，一些 ML 方法可以学到工程师的优化经验。最后，ML 方法在预测时通常非常快，所以如果不是规模太大，ML 方法通常在运行速度上比传统方法会有数量级上的优势。

4. 代表性的 ML for EDA 研究工作

事实上 ML for EDA 的工作远远不止前文介绍的两个例子。而这些工作可以通过多种方法进行分类。

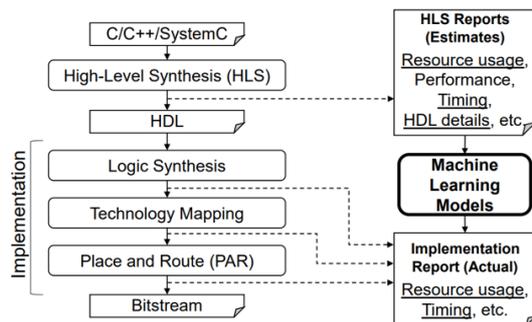
如果我们根据 ML 应用的阶段或步骤分类，ML 已经被研究者尝试用于绝大部分设计阶段。根据相关文章的总结 [1]，应用 ML 模型的阶段可以包括 1. 高阶综合 (HLS) 与设计空间搜索 (DSE)，2. 逻辑综合，3. 物理设计 (从布局规划到布线)，4. 光刻与制造，5. 验证与测试。除开数字电路之外，一些 ML 研究也对模拟电路的布局布线有一些尝试。



如果我们根据 ML 预测或优化的目标分类，ML 已经被研究者尝试用于大部分芯片优化目标，包括但不限于 1. 功耗 (power)，2. 频率与延迟 (delay or slack)，3. 面积与线长，4. 拥塞 (congestion) 与规则检查 (DRC)，5. 电压降 (IR drop) 与串扰 (crosstalk)，6. 可制造性 (Manufacturability)。

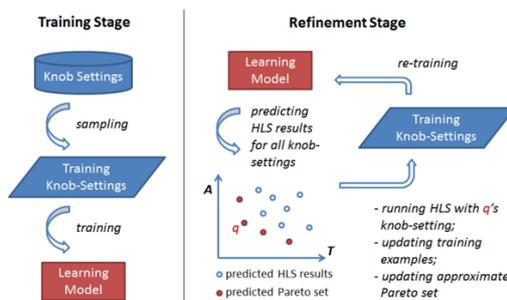
以下尝试简单列举一些具体的 ML for EDA 工作。篇幅所限，显然难以穷尽。因此同类型里尽量列举较早且较高引用的工作。对于更新更完整的 ML for EDA 工作总结，读者可以参考论文 [1] 或视频 [4]。

对于 FPGA 平台，在高阶综合 (HLS) 阶段，有工作训练 ML 模型来预测最终的 FPGA 资源的利用率以及是否时序收敛 [6]。其中预测的 FPGA 资源包括 RAM, FF, LUT 和 DSP 等。ML 模型是 MLP 或 XGBoost。ML 模型的输入，输出，使用阶段如下图所示。



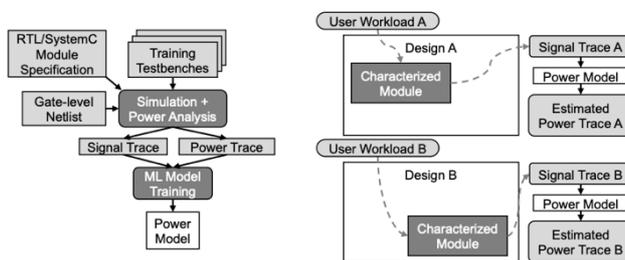
HLS 阶段对 FPGA 资源利用率进行预测 [6]

同样在 HLS 阶段，ML 可以进行设计空间探索 (DSE) [7]。基本的方法就是用 ML 模型的预测结果指导每一次采样，然后根据采样结果重新训练 ML 模型，反复循环直至采样的设计获得足够好的结果。这里用到的 ML 模型是随机森林 (RF)。这类探索任务也并不仅限于高阶综合。例如热门的用 ML 对 EDA 工具调参问题也属于相似的空间搜索任务。



用 ML (随机森林) 进行设计空间探索 (DSE) [7]

在 RTL 阶段，ML 可以用于对逻辑综合后的功耗进行快速估计 [8]。换句话说，ML 模型在设计早期进行快速功耗模拟。ML 模型的输入是芯片每个周期所有 RTL 信号的翻转值 (1 代表信号翻转，0 代表在该周期信号不变)，而输出是芯片每个周期的总功耗。选取这个输入是因为芯片的动态功耗与逻辑门的翻转率正相关。这个工作 [1] 尝试了包括线性回归，PCA 降维，CNN 的各类基础 ML 模型。

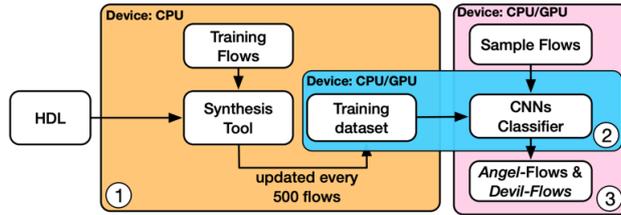


(a) Characterization phase

(b) Estimation phase

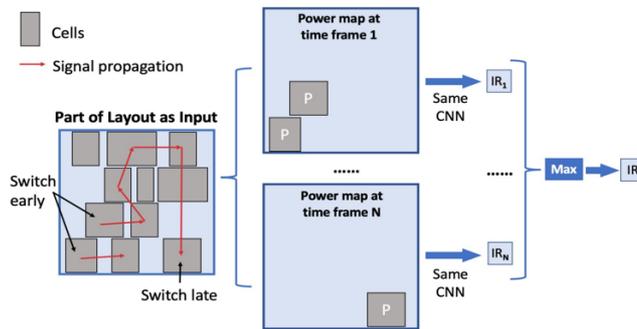
用 ML 模型 (CNN 或线性模型) 来模拟每周期的芯片功耗 [8]

在逻辑综合阶段，ML 模型可以用于对逻辑综合流程进行选择 [9]。这里 ML 模型被训练来预测最适合当前芯片设计的 synthesis transformation 的组合。该工作把各种 transformation 的组合方式编码成二维矩阵，然后用 CNN 模型来进行处理，预测它们是好流程还是坏流程。



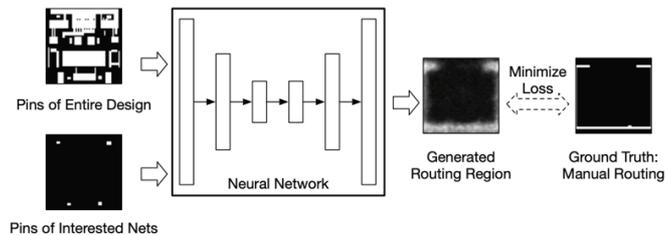
用 ML (CNN 模型) 选择合适的逻辑综合流程 [9]

在物理设计（从布局规划到布线）阶段，一开始提到的 macro 摆放和 DRC 热点预测两个例子都属于这个阶段的工作。另外一个例子是对芯片上的电压降（IR drop）分布进行预测 [10]。类似于之前的例子，CNN 模型可以输出电压降过高的热点（hotspot）。这里模型输入是瞬时的功耗分布。需要预测的电压降与区域内的瞬时的功耗需求成正相关。



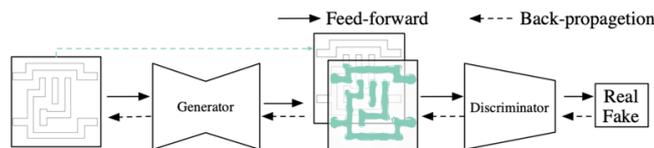
用 ML (CNN 模型) 预测电压降 (IR drop) 热点 [10]

除了以上这些用于数字电路的工作，对于模拟电路的探索也非常多。一个例子是直接使用 ML 生成对模拟电路绕线的指导（guidance） [11]。这里使用的 ML 模型是变分自编码器（VAE）。这个 VAE 模型的输入来源于模拟电路完成布局后的版图。



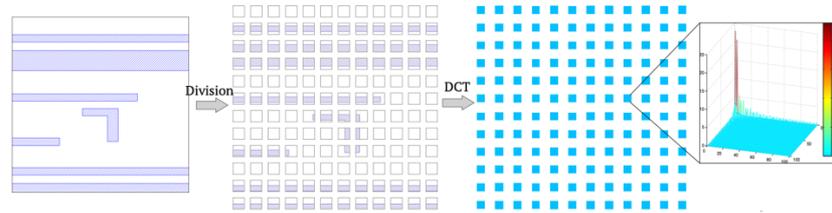
使用 ML (VAE 模型) 辅助模拟电路绕线 [11]

设计完成后，为了进行芯片制造，需要生成光刻机使用的掩膜（mask）。在掩膜生成的过程中，为了在晶圆上得到和设计相同的图形，光学邻近校正（OPC）是一个重要步骤。ML 可以用于进行 OPC [12]。这个工作使用生成对抗网络（GAN）来快速生成 OPC 后的掩膜。



使用 ML (GAN 模型) 进行光学邻近校正 (OPC) 过程 [12]

另外为了保证制造良率 (yield)，需要进行光刻 (lithography) 热点 (hotspot) 检测。光刻热点指的是制造后可能会出现短路或断路的区域。ML 可以用于预测对应的光刻热点 [13]。这个工作对每一块输入的芯片区域先进行离散余弦变换 (DCT)，然后使用 CNN 模型来预测该区域是否为光刻热点。



用 ML (CNN 模型) 对光刻 (lithography) 热点检测 [13]

5. 代表性的 ML for EDA 商业探索

而在工业界，EDA 公司和芯片设计公司也对 ML for EDA 的方向表达了很大兴趣。两大 EDA 公司 Cadence 和 Synopsys 已经推出了融合了一些 ML 算法的新产品。

例如 Cadence 公司的物理设计工具 Innovus 在新版本中已经里面集成了一些机器学习工具，比如说在布线 (routing) 前提供对 (布线后) 时序的更准确估计。这个 ML 功能似乎获得了一些不错的评价 [14]。另外之前有报道 Cadence 集成的 sign-off 工具 Project Virtus (Voltus + Tempus) 也使用了 ML [15]。更有影响力的是 Cadence 的 Cerebus 工具，通过基于强化学习的 ML 模型来调整整个 RTL-to-GDSII 的设计流程，减少对人工设计师的依赖 [16]。不过作为相关从业人员，我们很难了解这个产品的技术细节，例如这里的强化学习算法是如何迁移到不同的芯片设计的。

类似的，Synopsys 也在更早推出了 DSO.ai 工具，也是用强化学习的方法自动优化设计流程 [17]。Synopsys 的 PrimeTime ECO 工具也使用了 ML 模型。另外曾经的 Mentor Graphics (现 Siemens) 也推出过使用基于 ML 预测的光学邻近效应修正 (OPC) 工具。

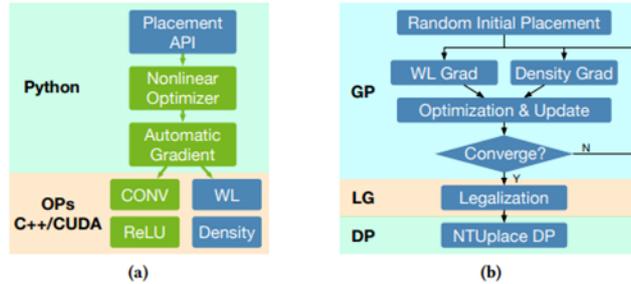
除了 EDA 公司之外，谷歌和英伟达的研究团队也积极探索了 ML for EDA 方向。这些是更偏研究的工作。具体可以参考谷歌 Jeff Dean 和英伟达 Bill Dally 的相关主题演讲。

而在国内，受美国半导体政策影响，近年来不少优秀的国产 EDA 初创公司涌现。一些公司也开始探索了 ML 在 EDA 领域的应用。例如某国内公司最近发布了基于机器学习技术与布局规划工具，应该是将类似前文提到的谷歌 macro 摆放算法[1]实现了产品落地。个人认为这是一个比较挑战性的任务。

6. 其他与 ML for EDA 相关的研究

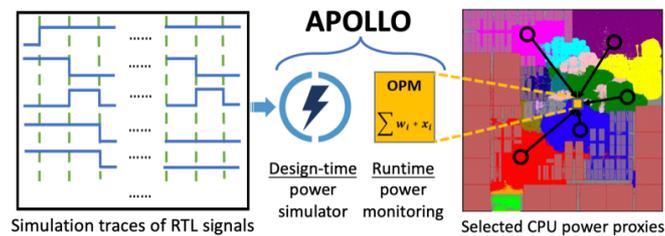
除去以上介绍到的工作，还有一些热点研究方向也经常被大家归类到 ML for EDA 的研究当中。但严格来说，个人认为这些方向也许不能完全属于狭义的 ML for EDA 的范畴。

使用 GPU 对 EDA 算法进行加速是近年来非常火热的研究方向。例如使用 GPU 加速布局过程 [18]。这些方法巧妙利用了 EDA 问题的优化过程与深度学习训练过程的相似性。于是它们可以利用已有的深度学习框架，例如 PyTorch，来进行基于 GPU 的快速优化。但这类方法不会去应用具体的 ML 模型，也没有基于数据的训练过程。



使用 GPU 加速芯片布局过程 [18]

另外一个常被归类为 ML for EDA 的方向是：在芯片上实现 ML 模型，将它用于进行芯片运行时的监测，控制，或管理。例如利用 ML 方法，开发一个片上的实时功耗检测工具 [19]。使用 ML 方法的目的是减少这个模型在芯片上的硬件开销。严格意义上说，这更接近芯片设计而非 EDA 工具范畴，也许可以将它归类到 ML for hardware design 的范畴。



使用 ML 开发的高效片上功耗检测模型 [19]

7. ML for EDA 面临的挑战

最后简单总结一下 ML for EDA 可能会面临的一些挑战。而针对这些挑战，近年来大家也进行了一些尝试。

第一，训练所需的芯片设计数据非常难以获得。对于这个问题，一方面一些研究者开始提供公开的芯片设计数据集 [20]。另一方面，研究者探索了在保护数据隐私条件下的模型训练，例如基于联邦学习的方法 [21]，或者对训练数据进行一些加密。

第二，开发和维护 ML 模型需要大量 ML 背景的工程师，而这对半导体公司来说并不容易。为了进一步提升自动化的程度，研究者探索了一些自动化的 ML 模型开发方法 [22]。这些工作一般基于深度学习领域已有的 AutoML 或者神经网络架构搜索 (NAS) 方法，然后对特定问题进行一些定制化设计。

第三，由于芯片设计与制程的差异，ML 方法的效果难以保证可迁移性。另外 ML 方法的安全可靠性也需要更多保证。已经有不少研究 ML 方法安全与可靠性的初步工作。论文 [23] 中提供了一些总结。

第四，从实际工程角度，ML 方法如何更好地融入现有的 EDA 工具与芯片设计流程。这一点也许需要研究者与工业界共同进行探索。

8. 结语

本文大致介绍了 ML for EDA 这一热门研究方向。可以看到，ML 方法在 EDA 领域提供的数据驱动的优势是极其独特的。考虑到这些特点，我个人确信更多的 ML 算法未来会融入芯片设计流程之中。但至于 ML 算法究竟会在未来的设计流程中占到多大比重，这当然取决于大家未来的探索与实践。非常期待 ML 方法能够推动 EDA 产业，特别是国产 EDA 产业的发展。

本文作者谢知遥，是香港科技大学电子与计算机工程系助理教授。他正在积极招收博士生，欢迎联系（个人主页：zhiyaoxie.com）。

参考文献

- [1] M. Rapp, et al. "MLCAD: A Survey of Research in Machine Learning for CAD", TCAD, 2022.
- [2] Z. Xie et al., "RouteNet: Routability prediction for mixed-size designs using convolutional neural network", ICCAD, 2018.
- [3] A. Mirhoseini, et al., "A graph placement methodology for fast chip design", Nature, 2021.
- [4] M. Ren, "Machine Learning for EDA Optimization", IEEE Seasonal School, [link] [youtube.com/watch?v=n5visiN1ncI&t=222s](https://www.youtube.com/watch?v=n5visiN1ncI&t=222s), 2022
- [5] Z. Xie, et al. "Intelligent Circuit Design and Implementation with Machine Learning", PhD Dissertation, 2022.
- [6] S. Dai, et al., "Fast and accurate estimation of quality of results in high-level synthesis with machine learning," FCCM, 2018.
- [7] H. Liu, et al., "On learning-based methods for design-space exploration with high-level synthesis", DAC, 2013.
- [8] Y. Zhou, et al., "PRIMAL: Power inference using machine learning." DAC, 2019.
- [9] C. Yu, et al., "Developing Synthesis Flows Without Human Knowledge." DAC, 2018.
- [10] Z. Xie, et al., "PowerNet: Transferable dynamic IR drop estimation via maximum convolutional neural network". ASP-DAC, 2020
- [11] K. Zhu, et al., "GeniusRoute: A New Analog Routing Paradigm Using Generative Neural Network Guidance." ICCAD, 2019.
- [12] H. Yang, et al., "GAN-OPC: Mask Optimization with Lithography-guided Generative Adversarial Nets." DAC, 2018.
- [13] H. Yang, et al., "Layout Hotspot Detection with Feature Tensor Generation and Deep Biased Learning." DAC, 2017.
- [14] "New CDNS 19.1 flow." DeepChip website, 2019. [link] www.deepchip.com/items/0587-05.html
- [15] "A True Signoff for 7nm and Beyond." SemiWiki website, 2018. [link] semiwiki.com/eda/cadence/7655-a-true-signoff-for-7nm-and-beyond/
- [16] "Cadence Cerebrus INTELLIGENT CHIP EXPLORER." Cadence website, 2022. [link] www.cadence.com/en_US/home/tools/digital-design-and-signoff/soc-implementation-and-floorplanning/cerebrus-intelligent-chip-explorer.html
- [17] "DSO.ai Achieve PPA Targets Faster with the World's First AI Application for Chip Design." Synopsys website, 2022. [link] www.synopsys.com/implementation-and-signoff/ml-ai-design/dso-ai.html
- [18] Y. Lin, et al. "DREAMPlace: Deep Learning Toolkit-Enabled GPU Acceleration for Modern VLSI Placement", DAC, 2021.
- [19] Z. Xie, et al., "APOLLO: An Automated Power Modeling Framework for Runtime Power Introspection in High-Volume Commercial Microprocessors," MICRO, 2021.
- [20] Z. Chai, et al., "CircuitNet: An open-source dataset for machine learning applications in electronic design automation," SCIS, 2022.
- [21] J. Pan, et al., "Towards Collaborative Intelligence: Routability Estimation based on Decentralized Private Data," DAC, 2022.
- [22] C. Chang, et al., "Automatic Routability Predictor Development Using Neural Architecture Search," ICCAD, 2021.
- [23] Z. Xie, et al., "The Dark Side: Security and Reliability Concerns in Machine Learning for EDA". TCAD, 2022.