

AI浪潮下的信息技术光电融合

张新全^{1,2*}, 余少华³

(1. 光通信技术和网络全国重点实验室, 武汉 430074; 2. 中国信息通信科技集团有限公司, 武汉 430074;
3. 鹏城实验室, 广东深圳 518000)

摘要: 人工智能 (AI) 有望催生第四次工业革命, 凸显我国信息技术自主发展的紧迫性和重要性。光电融合为 AI 发展提升硬件支撑能力, 是领域内需要重点发展的新质生产力技术。本文分析了光电融合的重要意义, 总结了光电融合的内涵、演进路径, 并介绍了相关技术的研究进展, 剖析了 AI 浪潮下信息技术光电融合发展面临的机遇和挑战。光电融合是 AI 基础设施发展的内在要求, AI 赋能可有效促进光电融合技术进步与突破。研究建议, 善用 AI 促进光电融合技术攻关、实施光电融合重大创新工程、营造充满活力的光电融合创新生态、培育和壮大具有国际竞争力的光电融合创新主体, 多措并举协同发力, 把握住 AI 浪潮下信息硬件技术变轨机遇, 有效提升我国信息科技自主发展能力。

关键词: 光电融合; 人工智能; 光互连; 异质异构集成

中图分类号: TN4 **文献标识码:** A

AI-Driven Photonics-Electronics Convergence for Information Technology

Zhang Xinquan^{1,2*}, Yu Shaohua³

(1. State Key Laboratory of Optical Communication Technologies and Networks, Wuhan 430074, China; 2. China Information Communication Technologies Group Corporation, Wuhan 430074, China; 3. Pengcheng Laboratory, Shenzhen 518000, Guangdong, China)

Abstract: Artificial intelligence (AI) is poised to usher in the fourth industrial revolution, highlighting the urgency and importance of China's independent development in information technology. Photonics-electronics convergence (PEC) offers critical hardware support for AI development and positions itself as a key technology for new quality productive forces. This study explores the significance, definition, and evolution path of PEC. It also presents the latest research advancements in related technologies and delves into the opportunities and challenges faced by PEC amidst the AI wave. As an inherent requirement for AI infrastructure development, PEC can be effectively promoted by AI. The study recommends leveraging AI to promote key technological breakthroughs in PEC research, implementing major innovation projects on PEC, fostering a vibrant innovation ecosystem, and nurturing globally competitive entities in this field. These measures aim to help China seize the opportunity of information hardware evolution amidst the AI wave and bolster the country's independent development capabilities in information technology.

Keywords: photonics-electronics convergence; artificial intelligence; optical interconnect; heterogeneous integration

收稿日期: 2025-11-14; 修回日期: 2026-03-12

通讯作者: *张新全, 光通信技术和网络全国重点实验室正高级工程师, 研究方向为信息光电子; E-mail: zhangxq@cict.com

资助项目: 中国工程院咨询项目“中国电子与信息工程科技前瞻储备性咨询研究”(2023-JB-14)

本刊网址: sscae.engineering.org.cn

一、前言

在机器学习 / 深度学习、大语言模型等技术进步掀起的人工智能 (AI) 热潮下, 信息化从数字化、网络化演进到智能化阶段^[1]。AI 有望催生第四次工业革命, 正在向各领域快速渗透, 深刻影响着技术发展和应用创新。2024 年的诺贝尔物理学奖授予了 AI 领域科学家, AI 成为近几年全球相关展会和论坛的关注焦点。在 2025 年召开的美国光纤通信会议上, 从运营商到设备商、器件商, 从网络架构到传送系统、可插拔光模块、关键器件, 各主体、各层次都在围绕 AI 进行部署和创新^[2]。相关研究指出, 过去 8 年间 AI 计算需求年均增长 10 倍, 下一个千倍级的性能提升靠什么? 仅对现有技术进行渐进式优化是远远不够的^[2]。

受制于物理特性, 当下电子信息技术面临多种技术壁垒, 在速率、带宽、能耗等维度接近“瓶颈”^[3]。电子芯片受到短沟道效应、量子隧穿效应等的制约, 制程升级步履维艰, 信息系统性能已持续六十余年的高速提升, 未来靠什么理论、什么技术延续? 变轨渐成信息技术继续进步的业界共识^[4]。

通过光电融合充分发挥出光子、电子各自的性能优势来推动信息系统在速率、带宽、能效等方面继续大幅、快速进步, 是业界较为认可的信息技术变轨方案^[5]。需要指出的是, 尽管变轨在学术上还有多种可能方案 (采用新的原理、结构和材料, 如: 自旋器件、量子元胞自动机、碳纳米管、磁通量器件等), 但从可实现性、产业基础等多角度考虑, 产业界目前的积极探索和实践多集中于光电融合方案, 正在对光电共封、光输入 / 输出 (I/O)、光电集成、光电背板、光子 AI 加速、光电计算等方面进行攻关研究^[6], 光电融合已成为领域内重点发展的新质生产力技术^[7]。

本文对 AI 浪潮下光电融合的价值、内涵、发展路径进行了探讨, 总结了相关技术进展; 阐述了光电融合对于 AI 发展的重要支撑作用, 以及 AI 对光电融合技术进步的助力和促进作用, 两者相互推动, 共同为信息技术进步开辟空间; 从善用 AI 促进光电融合技术攻关、实施光电融合重大创新工程、营造充满活力的光电融合创新生态、培育壮大有国际竞争力的光电融合创新主体等方面提出对策建议, 谋划通过光电融合提升我国 AI 和信息技术的自主发展能力。

二、光电融合的内涵及研究进展

(一) 光电融合的内涵

光电融合是强调光子和电子的融合发展, 通过两者的器件一体化、功能融合化来共同解决信息技术面临的迫切挑战^[6], 以适应 AI、数字孪生、元宇宙等应用对速率、带宽、能效等的不断攀升需求。光电融合是为了更好发挥光子、电子的各自优势, 并规避各自不足。光电融合的对象是电子和光子, 两种粒子对材料与工艺的要求有较大差异, 其融合难度高, 不能简单等同于电子集成的自然延伸^[8]。

光电融合主要在芯片和系统两个层面展开。芯片层面的光电融合最具紧迫性, 不仅可以有效解决微距高速互联的时延和功耗问题, 而且可以推动芯片从单一功能向多功能乃至智能化演进, 并向新的应用 (如微波光子、光子神经网络、光量子信息等) 快速渗透和裂变^[6]。光电单片集成是融合的较理想途径, 但近期光电混合集成的可实现性更强, 业界探索亦更为积极, 如用于芯片级光互连的共封装光学 (CPO)、光 I/O 等技术。

系统层面的光电融合是基于光电融合来重构信息系统。光电融合不是光子、电子技术的简单叠加、简单协同, 而是两者有机结合、深度一体化, 必然引发信息组件、设备、系统的重构^[6]。当前对系统层面光电融合的关注不多, 但其意义已初显, 如解聚计算、资源池化、光电混合计算等^[9]。

光电融合将促使光子技术在信息领域的应用范围不断扩大、作用不断提升。一方面, 光子由传输技术泛化为信息与通信技术 (ICT) 全域的泛在连接技术, 不仅从地面进入海洋、太空的宏尺度, 而且从微尺度上进入架间、板间、板内、封装内、片内; 另一方面, 光子由带宽提供技术泛化为 ICT 全域的硬件基础技术, 从传输和连接进入计算、处理、路由等复杂功能域^[6]。

光子要泛化为与电子相当的基础性信息硬件技术, 光子器件水平须有大进步。数十年来微电子技术的发展, 使得电子器件不仅能大规模生产, 而且性能、可靠性不断提升, 成本、体积不断缩减, 从而满足现代信息系统的要求。因此, 基于电子器件的历史成功经验, 以及近十年光子器件的发展态势, 我们提出光电子发展的“微电子化”^[6]。

光电子微电子化 (见图 1), 强调以集成为轨道

朝着光电融合方向发展，通过集成使成本、性能、体积、能效等按照“类摩尔定律”推动信息系统继续向前演进，延续技术进步与应用规模提升的良性互促^[6]。

光电子的微电子化发展，主要呈现3个特征：一是芯片平台硅基化，二是集成规模稳步提升，三是制造模式向无晶圆厂模式（fabless）演进^[6]。芯片平台硅基化，在光电子集成必然是多材料异质集成的前提下，业界逐渐将硅作为基础平台。制造fabless化，是光电子集成的制造与设计分离，如硅光子制造由格罗方德半导体股份有限公司、台湾积体电路制造股份有限公司等晶圆厂完成。

光电子器件集成化并日益向光电融合发展，是支撑AI继续发展并最终可能形成电子、光子并重的信息硬件格局，从电子信息时代迁移到光电信息时代^[5]。但光电融合需要经过一个持续演进的长期过程（见图2），经过不断地创新突破方能使信息技术从“电子为主、光子为辅”的“电算光传”离散分工迈向“电子、光子并重”的光电深度融合。

（二）光电融合的研究进展

光电融合在近几年得到业界较多关注，当前的研究重点是将光互连引入信息设备内部、深部，通过光收发组件芯片化（实现组件内光、电各功能集

成）和布线光化（光背板、光印制线路板、光中介板等）代替板间、板内、片间、片内电子已力有不逮的高速连接，从而使信息设备在功能架构、资源组织等方面不再受限于电子连接的能耗、带宽、距离限制，有效满足当下AI浪潮对信息硬件技术的性能需求。远期，光计算等光信息处理技术突破将进一步拓展光电融合的广度和深度。

1. 单片光电集成

单片光电集成是在单个芯片衬底上将光、电器件集成在一起。尽管Somekh S.和Yarive A.在1972年即已提出该设想，但直到硅光子能够实现大部分光器件后，业界才开始真正推进硅基的单片光电集成^[10]。

在光收发器方面，国际商业机器公司（IBM）早在2016年就采用55 nm互补金属氧化物半导体（CMOS）工艺实现了25 Gb/s光收发器（激光器是片外）的单片光电集成^[11]。Sicoya公司则采用双极型-互补金属氧化物半导体（BiCMOS）工艺研制出100 G光收发芯片，其中光器件（包括光栅耦合器、光探测器、光调制器、光波导等）制作于芯片左侧的绝缘体上硅（SOI）区域，电器件（驱动器、时钟恢复单元、跨阻放大器、控制器等）制作于芯片右侧的体硅区域^[12]。2021年，苏黎世联邦理工大学实现了功耗低至1.91 pJ/bit的单片集成10 Gbps光收发芯片（尺寸为0.015 mm²，45 nm工艺）^[13]。同样采用45 nm工艺，2023年Luminous Computing的研究团队将16个通道的调制器、驱动器和控制电路集成到面积为0.38 mm²的芯片上，实现了功耗仅约为7.6 mW的1.8 T发射机^[14]。2024年，Ayar Labs在国际固态电路会议（ISSCC）上展示了名为Shasta的光I/O芯片，共有8个端口，每个端口有8个波长的双向通道（每通道速率为32 Gbps）^[15]。该芯片基于格罗方德半导体股份有限公司的45 nm SOI工艺，在300 mm平台上实现了硅基光器件和电器件的单片集成，采用半脊调制器，通过V形槽边缘耦合实现自对齐，从而简化光纤连接，发送能效为1.87 pJ/bit，接收能效为2.09 pJ/bit。中国科学院半导体研究所在2024年研制出单向带宽为256 Gbps（4个波长通道，每个波长通道为64 Gbps）的单片集成光互连芯片（45 nm工艺，包括微环调制器、微环滤波器、热相移器、光电探测器等光学器件，以及驱动器、跨阻放大器等电学器件），发送能效

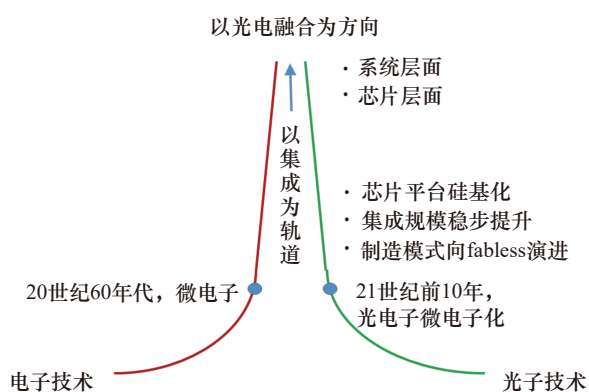


图1 光电子发展的微电子化



图2 信息技术从“电算光传”的离散分工迈向光电深度融合

为 1.6 pJ/bit，接收能效为 1.25 pJ/bit^[16]。

在处理器方面，2017 年加利福尼亚大学伯克利分校和麻省理工学院采用 45 nm SOI CMOS 工艺制作出单片集成的处理器系统（共约 7×10^7 个晶体管 and 850 个光子元件，实现微处理器与存储器之间的片上光互连，并在次年改用体硅 CMOS 工艺以优化成本^[6,17]。

芯片层面的光电融合，单片光电集成是较理想的方式^[6]。相比于基于分立芯片的封装集成，单片集成有如下优势：一是尺寸小；二是避免封装键合所引入寄生效应对性能的劣化；三是避免高频信号进出芯片导致的功耗；四是降低了封装复杂度。但是，由于单片光电集成的技术挑战大（光、电在材料与工艺上有不能忽视的较大差异），从实际的综合效果看至少目前尚不是最佳选择，所以单片光电集成近几年的关注度有所降低。

2. 混合光电集成

混合光电集成属于超越摩尔（MtM）路径的技术^[18]。它将光、电芯片通过 2D、2.5D、3D 等先进封装技术集成一个整体^[19]。

2D 方式工艺简单，但尺寸相对较大。它是将电、光芯片水平排列在封装基板上，采用引线键合进行内部芯片互连。微电子研究中心将采用 55 nm BiCMOS 工艺制作的 Driver、TIA 等电芯片与硅光芯片进行 2D 集成，实现 200 Gbps 的收发（4 个波长通道，每个波长通道为 50 Gbps），能效为 4.2 pJ/bit^[20]。

2.5D 方式在当前被认为是成本、性能、成熟度、工艺难度等较均衡的方案。它将电、光芯片以覆晶封装方式安装到中介板上，通过中介板上的金属布线实现内部芯片互连。英特尔公司以 2.5D 方式实现了光发射机、光接收机的光电集成，相关电芯片采用 28 nm 工艺制程。其中，256 Gbps 速率的光接收机（8 个波长通道，单波长通道为 32 Gbps）的能效为 3.8 pJ/bit^[21]；400 Gbps 速率光发射机（8 个波长通道，每个波长通道为 50 Gbps）的能效为 2.5 pJ/bit，所采用的硅基微环调制器驱动电压仅为 0.4 V（能效约为 83 aJ/bit）^[22]。

3D 方式的工艺要求最高。它将光、电芯片垂直堆叠，由于内部互连长度极短，尤其适合高速、高频。哥伦比亚大学的研究人员把电芯片覆晶封装在光芯片上，以 3D 方式实现了 64 通道硅光收发器集成^[23]，光、电芯片以 25 μm 铜柱进行电学互连，

带宽密度达到了 5 Tbit/s/mm²。

3. 硅基光电子

硅基光电子是利用硅材料及其兼容的制造工艺，通过与微电子工艺相匹配的方法来制造光子芯片，实现光子与电子器件的紧密结合^[24]。在英特尔公司实现硅基调制器的关键突破^[25]后，硅光能制作除光源、光放大以外绝大部分有源与无源器件，不仅在体积、成本、产业化条件等方面相对于其他材料平台有明显优势，而且因为与电子器件集成平台同材料而尤其利于光电融合。硅光单器件的性能水平不断提升，硅基慢光调制器^[26]、硅基薄膜铌酸锂调制器^[27]等相继突破 110 GHz 带宽。另外，硅光的集成规模也在快速增长，已能实现超万个器件的集成^[28]。众多器件集成于硅光芯片上，不仅体积小，而且可靠性得到了提升（如 100 G 光收发器的缺陷已仅为 30 个/百万^[29]）。因此，硅光成为超太比特每秒速率后光模块的主要技术选择。美国海军研究实验室指出，硅基光电子技术是集成电路的下一代革命，将之列入未来 25 年的 25 项关键技术之一^[30]。

硅是间接带隙材料，发光效率低（硅的光发射是声子辅助低概率过程，复合寿命仅毫秒量级），直接制备激光器一直未能有效突破^[31]。产业界目前多在硅光子芯片的片外提供光源，或以覆晶封装方式将硅光子芯片与光源芯片进行混合集成^[32]。英特尔公司认为异质集成方式更具优势^[33]，其与 CEA-Leti、Skorpios 等将外延键合到硅基上，再采用半导体精密工艺制作片上激光器^[34]。日本电报电话公司开发出超薄 III-V 族激光器（厚度 $< 1 \mu\text{m}$ ，功耗较普通垂直腔面发射激光器（VCSEL）降低 70%），通过水平电流设计实现高效激光发射，并采用微间隙对准技术实现光芯片与硅基电芯片的亚微米级精准对位（误差 $< 0.1 \mu\text{m}$ ）^[35]。最理想的方式是在硅衬底上进行外延生长来直接制备片上激光器，不过，由于在非极性硅衬底上生长极性化合物会导致反相畴^[6]，且二者晶格失配严重、热膨胀系数差异大，业界尝试通过引入多个厚度、多种材料的缓冲层来弱化这些不利影响^[36]，相关技术已取得较大进步但离商用还有一定距离。为此，研究人员提出通过深窄沟槽限制适配缺陷并在沟槽上端形成 III-V 材料层状脊柱以制作量子阱激光器的方法^[37]。该沟槽（宽 120 nm，深 300 nm）在二氧化硅层，沟槽底部是硅，III-V 材料沉积于沟槽中，由于源自硅与 III-

V 材料界面的缺陷会终止于沟槽顶端, 使得沟槽高度以上能够生长出无缺陷的外延结构。

4. 芯片级光互连

芯片级光互连是将光收发芯片与大型电处理芯片集成在一起, 以显著提升片间互联的速率、容量、距离, 并显著减少延迟和功耗。芯片级光互连的研究, 目前主要集中在 CPO 和光 I/O 两个方向。

CPO 是将大型交换芯片与光芯片封装在一起, 与外部其他芯片形成光互连, 现阶段主要针对以太网交换机。CPO 是可插拔光模块的替代, 可规避交换芯片至面板的较长距离高速 (超 100 Gbps) 电互连导致的高功耗、大时延, 并规避容量几十上百太比特每秒后交换机面板的端口密度问题, 使带宽密度得到数量级提升。CPO 所采用的光收发芯片, 在串行器速率达到 224 Gbps 或以上、单通道波特率超 56 G 后, 硅光子方案较 VCSEL 阵列方案得到更多认可, 因为硅光子芯片无需气密封装, 关键器件 (硅光调制器和探测器) 均有能力支持 56 G 以上波特率, 且易与相关电芯片集成, 在集成度和规模产业化方面优势较明显。2024 年, 横浜国立大学将硅慢光马赫-曾德尔调制器 (MZ) 调制芯片和电驱动芯片 (基于 IHP 的 130-nm SiGe BiCMOS 工艺) 通过铝线键合, 功耗低至 1.9 pJ/bit^[38]。2024 年 5 月, 国家信息光子创新中心 (NOEIC) 和鹏城实验室联合团队完成 2 Tbps 硅光互连芯粒的研制和功能验证, 采用三维堆叠封装工艺实现了 3D 硅基光电芯粒集成, 单片的单向互连带宽为 (8×256) Gbps^[39]。当然, 因为 VCSEL 阵列方案确有成本和功耗优势, 所以其研究也在继续中。在 2024 年欧洲光通信会议上, 富士通株式会社展示了基于 1060 nm 单模 VCSEL 阵列和单模多芯光纤 (MCF) 的超紧凑 CPO 收发器, 其正面采用垂直发射与单模 MCF 直连, 背面通过电芯片与基板相连, 以提升紧凑性^[40]。

光 I/O 强调将计算芯片与光芯片集成在一起, 与其他计算芯片、存储芯片等形成光互连。光 I/O 聚焦于计算领域, 较之于电 I/O 可以显著提升算力构建的规模、效率。光 I/O 对集成度、规模化制造的要求较 CPO 更高, 目前光 I/O 芯片技术方案大多都是基于硅光子的。2023 年, 康奈尔大学发表了带宽密度为 5.3 Tbps/mm、能效为 120 fJ/bit 的光 I/O, 共 80 个 10 Gbps 通道^[41]。其中, 光芯片采用 AIM Photonics 的 12 in (1 in=2.54 cm) 硅光工艺, 电芯

片采用 TSMC 的 28 nm CMOS 工艺, 两者高密度倒装键合 (间距仅为 15 μm); 采用驱动电压仅 1 V 的垂直 PN 结微盘调制器, 使发射功耗低至 50 fJ/bit。英特尔公司和 Ayar Labs 联合实验中的光 I/O 总带宽可达 40 Tb/s, 带宽密度约为 5 Tbps/mm, 能效约为 3 pJ/bit, 时延约为 5 ns^[42]。Ayar Labs 还展示了其匹配通用芯粒互连技术标准 (UCIe) 电气接口的光 I/O 芯片。在 2024 年开放计算中国峰会上, Celestial AI 展示了其速率为 14.4 Tbps, 功耗为 2.4 pJ/bit (另有外置光源为 0.7 pJ/bit) 的光 I/O 单芯片方案^[43]。该方案一是利用锗硅电吸收调制器 (GeSi EAM) 对温度的不敏感来规避百瓦级 xPU 引发的热稳定性问题; 二是配套电芯片以 2.5D 或 3D 方式与光芯片进行高密度封装以实现无数字信号处理器 (DSP) 的直驱; 三是通过电-光链路协同优化以去掉用于速率适配调整的变速箱, 从而降低功耗优化时延。

业界龙头在以芯片级光互连促进算力和其他信息硬件发展上基本达成共识。英特尔公司则在 2023 年 8 月发布了采用光 I/O (1 Tb/s) 实现的 8 核 528 线程可编程集成统一存储架构 (PIUMA) 处理器^[44]。英伟达公司已将 CPO 引入其新一代的 IB 交换机, 并把光 I/O 列入其 NVlink 演进路线图; 博通公司 (Broadcom) 近两年连续发布了支持 CPO 的 25.6 T、51.2 T 交换引擎芯片。

5. 光电计算

光计算有数字和模拟两种方式。数字光计算可进行通用计算, 基础是光子逻辑门 (光子多稳态开关器件)。由于光子逻辑门的实现难度极高, 导致数字光计算长期无法取得实用化进展^[5]。模拟光计算是利用光子器件特性实现特定计算功能, 相对而言较简单, 属于专用计算, 近几年不断取得研究进展^[45,46], 在 AI 加速、伊辛问题求解等方面优势明显^[47]。例如, 上海曦智科技股份有限公司 (Lightelligence) 基于硅光技术制作的光子算术引擎 Pace, 在伊辛问题求解上比电子图形处理器 (GPU) 方案快百倍^[48]。2024 年, 清华大学的研究团队提出分布式广度光计算架构, 并研制出光计算芯片“太极-I” (可重构、高并行的浅层神经网络), 能效达到 160 TOPS/W^[49]; 2025 年, 该团队进一步提出全前向智能光计算训练架构 (利用空间对称性和洛伦兹互易性消除梯度下降训练中的反向传播必要性), 并将之应用到芯片“太极-II”, 可使百

万参数的神经网络训练速度提升一个数量级，代表性智能分类任务准确率提升40%^[50]。

光电融合计算为计算硬件实现提供了新路径。电子计算灵活、成熟、易数字化；光子计算能耗低、高速、可并行^[51]。由于二者在计算上各有长短，有望构建出二者功能协作、性能互补的大算力、低功耗的光电融合计算系统。

三、AI 浪潮下光电融合面临的机遇与挑战

(一) AI 浪潮凸显光电融合的重要意义

AI 技术在近十年取得显著进步。AlphaGo 围棋水平超过人类顶尖棋手，机器学习 / 深度学习广泛应用到视听处理、图像识别、物体检测等方面，强化学习推动机器人等各类无人系统性能跃升，Transformer 架构使大模型展示出文本、图像、视频等各种模态的强大生成能力。

AI 性能高度依赖算力，其能力提升是以算力的指数级增长为支撑的。大模型需处理海量多维信息，进行数十亿、数百亿参数规模的训练，算力需求增长远超电子计算能力增长。电子计算存在“不可能三角”，模型参数规模、实时计算效率、能效三者因相互制约而不能同时最优，导致算力成为当前 AI 发展的重要限制因素。电子逻辑开关的实际能耗高于朗道尔 (Landauer) 公式理论值的数个数量级，如基于 3 nm 制程鳍式场效应晶体管 (FinFET) 的 CMOS 反相器翻转需耗能约 0.039 fJ。以英伟达公司 B200 芯片为例，它执行 16 位浮点数据运算的速度为 2250 Tera FLOPS，功耗近千瓦，能效为 400 pJ/op^[52]。训练 1750 亿参数的 GPT-3 模型，处理 1000 个词元需执行 87.5 T 次的 MAC 操作，训练能耗高达 36 MW (以每个 MAC 操作耗能 0.25 fJ 来计算)^[52]。

AI 对算力的极高要求，需要计算硬件底层范式的创新，光电融合是重要的创新路径。虽然可以通过算法、架构的优化来缓解 AI 对算力的消耗，但 AI 的大规模应用必然使未来的算力规模持续攀升。光子计算 (如光神经网络) 具有并行运算、能效高、计算速度快等特点。目前，光神经网络主要有两种实现方案，一是采用衍射光学分立元件^[53]、超表面^[54]等的自由空间方案，有利于发挥三维光学互连优势进行调控；二是采用基于片上衍射^[55]、马赫-

曾德尔干涉仪 (MZI) 阵列^[56]等的集成方案，具有体积小、易扩展等特点。此外，AI 系统的高复杂性要求神经网络具有高灵活性，通过采用可调谐材料 (如 Ge₂Sb₂Te₅^[57]、液晶^[58]等) 和器件^[59]，结合原位学习策略^[60]等，可重构神经网络，能有效实现灵活的训练加速。当前，光神经网络、光子张量处理等各种光子计算技术在特定任务上已展示出远超电子计算 (如 GPU) 的速度和能效 (速度提升 3 个数量级，能效提升 6 个数量级)^[61-63]。但是，受制于实际器件的非理想性以及其它因素，光神经网络在准确性和泛化能力方面逊于电子神经网络^[67]。因此，光电深度融合的计算硬件才能发挥各自所长，成为 AI 算力的突破性解决方案，为 AI 发展提供颠覆性硬件支撑^[64,65]。

通过光电融合在片间 (电处理器芯片之间，电处理器芯片与存储器芯片之间) 引入光互连，也是提升 AI 算力的一个重要方面。由于算力横向扩展方式要求的速率距离积超过电互连技术的极限，智算中心的跨机柜互连已完全依赖光，当前 800 Gbps 可插拔光模块已大规模采用，1.6 Tbps 可插拔光模块也部分采用。对于算力的垂直扩展方式，目前大都在一个机柜内部署，由于成本约束且距离有限，仍采用电互连。但随着 GPU 对带宽的要求攀升，英伟达公司已计划在其下一代架构 (Rubin) 中将光互连引入机柜内。片间光互连在能效、带宽等方面的优势，使其成为算力垂直扩展的必然选择，CPO、光 I/O 等技术近年来进步迅速。表 1 中列出了垂直扩展对光互连的性能要求^[2]，博通公司认为，2027 年光互连技术可能会满足垂直扩展通信要求^[2]，这个判断与英伟达公司的产品路线图基本吻合。因此，光互连必然进入 AI 服务器内部，与各类电处理芯片深度融合。

表 1 垂直扩展对光互连的性能要求

维度	单位	垂直扩展 通信	封装外存储 接口	综合 要求
片间距离	m	>20	>1	>20
带宽	Tbps/dir	>50	20~50+	>100+
往返时延	ns	<1000	<500	<500
带宽密度	Tbps/mm	>4	>4	>4
能效	pJ/b	<4	<4	<4
纠前误码率		10 ⁻⁹	10 ⁻¹²	10 ⁻¹²

此外,光电融合交换是AI智算中心网络的发展趋势。尽管智算中心的网络连接媒质已全部是光纤,但网络交换仍然依赖电交换。目前,谷歌公司已将光交换引入其智算中心网络中作为顶层交换技术^[66],以缓解成本、能耗压力。在算力不断扩张的趋势下,智算中心网络的规模还将继续扩展,而网络的成本、能耗已经成为智算中心发展的限制因素之一,相干光系统与芯片间的光/电互连一起形成智算集群网络层面的光电融合解决方案,光电融合交换、光电混合组网是智算网络向更高性能、更低成本、更高能效不断演进的必然选择^[67]。

可见,光电深度融合是AI基础设施发展的内在要求,光电融合技术的突破将有效促进AI技术进步和应用推广,具有重要意义。

(二) AI赋能光电融合

光电融合必然带来新的技术挑战。以CPO、芯粒等光电融合集成为例,多种光、电器件和芯粒通过微凸块、TSV/TGV、中介板等各种2D/2.5D/3D先进封装高密度集成在一起,由于空间极狭小、信号频率/速率极高,叠加多种材料、多种功能,导致呈现强烈的“电-光-热-力”多物理场交互效应,电磁环境极为复杂,信号串扰和噪声的分析与应对要求极高,热管理极困难(如局部温升和大尺度温度梯度会诱发界面剥离、裂纹扩展等热致机械失效,甚至影响材料的物理特性)^[68]。这些困难不仅直接影响信号的完整性、热可靠性、电磁兼容性,还决定着器件性能、功能的长期可靠性。

因此,光电融合芯片需要进行“电-光-热-力”等多物理场建模与分析、仿真,进而开展协同优化的设计,而AI技术是实现的关键手段。AI在复杂模式识别和大规模数据处理方面能力优异,深度学习和强化学习对微纳光电器件设计和优化具有重要意义,善于从庞大且复杂的数据集中提取有效信息,识别出潜在的规律和模式,从而快速、准确地预测器件性能或进行逆向设计,大幅减少迭代仿真所需的时间和计算资源,并改善设计效率和性能^[69]。2025年,普渡大学的团队采用基于图像的机器学习构建异构芯粒多物理场预测模型,能够实时预测封装内的热分布图,并结合动态秩揭示优化算法直接在物理域上进行布局优化^[70]。通过将“电-热-力”约束直接嵌入优化过程,实现多物理场性

能与布局的同步迭代,在保证准确性的同时大幅降低传统有限元仿真的计算开销,显著提升了设计效率与系统性能。

将AI应用到光电融合硬件的正向建模与逆向设计过程中,可显著提高问题解决效率。在正向建模中,通过从数据分布中学习和提取特征,神经网络为光电参数和物理响应之间建立映射关系,深度学习比传统方法的速度更快、精度更高;在逆向设计中,AI根据目标性能给出光电参数,为光电融合器件结构设计、材料设计提供优化方案^[71]。通过AI提升光电器件联合优化水平,以直调激光器(DML)为例,DML有低成本的优势,但超百吉波特率后性能不理想。通过AI建立可解释的激光器模型,在AI辅助下将DML参数和电数字信号处理器参数进行联合优化,以实现端到端超高速率^[72]。

由于光电融合芯片的设计方法学近似于泛模拟电芯片设计方法学,AI能够显著改善自动化设计水平和效率。深度学习和仿真结合,可以高效实现器件参数与版图布局的自动、最佳优化^[24]。2025年,光通信技术和网络全国重点实验室等单位基于GPU实现超紧凑可扩展模分复用器的逆向设计,器件尺寸均为数微米量级,精度在纳米级别,插损和串扰特性优良^[73]。传统光电器件往往依赖人为参数调节与结构直觉,在器件紧凑化与高阶模扩展方面受限明显,基于GPU加速的全波伴随拓扑优化方法可将器件设计问题转化为连续介质分布优化,在微米级设计区域内自动生成复杂的光场分布结构,实现功能最优匹配。该设计框架具备可拓展性与模块化特征,具备自动化、快速、物理约束可解释等优势,与机器学习等技术结合可作为AI辅助光电器件反演设计的可行路径,是实现硅基光电融合设计的重要工具^[74]。

此外,通过基于机器学习的制造工艺实现自优化和自适应,对于硅基光电器件集成制造和封测也有重要意义^[24]。

可见,AI能够赋能光电融合研究,有利于推动光电融合设计创新,并大幅提升仿真速度和精度。传统设计方法是基于物理定律和经验公式建模,反复仿真验证,而AI是直接基于大量实验数据通过深度学习来识别光电融合器件的设计规律、预测性能、优化路径。例如,谷歌公司2023年发布的材料探索图形网络(GNoME),将稳定材料种类从人

类已知的约4万种增加到约38万种^[75]。随着AI技术的不断进步，AI有望成为光电融合的内生要素，显著加速光电融合的创新突破^[69]。

（三）光电融合是我国在AI时代的变轨机遇

毋庸讳言，现代信息技术发轫于西方，我国通过制造优势和应用创新逐渐成长为全球信息产业的重要一极。芯片是AI等信息技术的基础硬件，被西方当做科技博弈的主要抓手。

一是通过关键芯片的“点穴”式禁运来摧毁中国信息企业的国际竞争力。中国主要的信息企业都被列入美国的商务部实体清单，芯片断供持续变本加厉。

二是通过长臂管辖等手段试图对中国信息产业釜底抽薪。美国通过组建“芯片四方联盟”(Chip 4)等对中国封禁先进制程产线，禁止3 nm及更先进的电子设计自动化工具出口中国，拉拢其他国家在2023年全面限制先进制程装备，试图断绝中国发展芯片先进制程的可能性。

三是通过“芯片法案”等举措进一步强化美国的主导能力。美国提出，到2030年，生产全球至少20%的先进逻辑芯片、建成至少三个大规模先进封装产线、新增四座成熟节点大型晶圆厂，确保掌控从石英到芯片的完整供应链^[76]。

光电融合有利于我国破除芯片上的被动，建立包括AI在内的信息产业自主发展能力。我国在芯片方面缺乏有实力的盟友，“自强”是唯一选项。光电融合是我们有能力抓住的信息硬件技术变轨机遇：一是我国在光子方面的科研、产业实力水平与西方基本相当；二是光子集成对先进装备的依赖较低（当前主流产线是45 nm、28 nm制程），不会完全受制于外。

（四）光电融合面临技术和国际竞争挑战

在技术维度，光电融合还有诸多壁垒需要创新突破。最近十年，业界在光互连、光子集成、光电共封、光电集成、光交换、微波光子、光计算等方面不断探索，取得了一些进展。但光电融合仍面临许多技术挑战：一方面，光子和电子的物理特性差异显著，导致融合存在许多技术困难^[4]；另一方面，人类目前对光子的操控、运用水平较显著地逊于电子，光子集成的技术水平和产业化水平也显著逊于

电子集成，面临材料兼容性、光场多维度调控、设计理论与工艺平台等多方面的挑战，制约了光子的潜力发挥^[77,78]。

在国际竞争维度，光电融合还需要面对外部博弈与竞争。光电融合与半导体材料、先进封装、设计软件等联系紧密，国外禁运增加了我们的发展难度。而且，美国、日本、欧洲等国家和地区也在积极部署光电融合的研究。日本设立多个相关项目，致力基于光电融合实现板级服务器和片上数据中心^[79]，并提出光电融合设备的四阶段发展计划^[80]。美国国防部高级研究计划局布局了一系列光电融合相关技术研究计划，如光电异质集成（EPHI）、封装内极致光互联（PIPES）、光电融合嵌入式微处理器（POEM）、异构自适应光子接口（HAPPI）等，2026年1月发布的面向可扩展系统PIC架构（PICASSO）项目则提出通过超大规模光子集成超越计算、AI加速、传感等电子系统。电气与电子工程师协会（IEEE）每两年发布或更新异质集成路线图（HIR），为行业提供明确的技术指引，对光电融合有较强促进作用^[81]。

四、AI时代推进光电融合的发展建议

AI正在广泛、深入地重塑社会生产生活的方方面面。通过光电融合为AI发展提升硬件条件，通过AI赋能推动光电融合技术突破，是推动我国信息科技自主发展的有效路径。为了促进我国更好地把握住光电融合机遇，提出以下发展建议。

（一）善用AI促进光电融合技术攻关

“AI科学家（AI for Science）”自主发现新材料、新结构、新机制的时刻正在临近，AI将为光电融合技术创新开辟颠覆性路径，我们需要高度重视并善于利用AI开展光电融合技术攻关。

第一步，发挥AI的赋能作用。光电融合涉及多物理场耦合（电、光、热、力等）、多材料体系协同（硅、III-V族、薄膜铌酸锂、相变材料等）、多尺度工艺整合，传统基于人类直觉与经验的设计已难以应对，AI赋能的逆向设计能以“数据-物理”双驱动方式适应这种复杂的高维约束设计，并通过双层优化框架确保工艺的可实现性。AI赋能“设计-工艺”协同优化和“系统-工艺”协同优化，

实现材料参数（折射率、损耗、非线性系数）、工艺变量（温度预算、薄膜厚度、退火条件）与器件性能（插入损耗、消光比、带宽）的跨材料体系快速性能预测与参数寻优。此外，AI在三维堆叠布局优化、制造工艺控制、晶圆测试和根因分析等方面也可发挥重要作用。

第二步，探索将AI变成光电融合技术创新的内生要素，从根本上重塑光电融合技术研发范式。AI不只是可作为“多物理场-多材料-多尺度”复杂场景下的方法或工具，而且是一种基于数据智能与物理原理协同演化的光电融合全链条创新范式。

（二）实施光电融合重大创新工程

将光电融合作为领域内重点发展的新质生产力技术，实施光电融合重大创新工程。新质生产力的要义，是以科技创新摆脱传统路径依赖。光电融合可以改变信息技术对传统电子路径的过度依赖，突破现有信息技术遭遇的诸多瓶颈制约。

实施光电融合重大创新工程，一是攻关光电融合的关键技术，如片上光源、2.5D/3D先进封装、异构集成、光I/O、光神经网络，探索新机理材料，探索多维材料结构，推动信息技术硬件继续向前演进，避免被“卡脖子”；二是建设光电融合的信息基础设施，通过光电融合的服务器、数据中心、网络等支撑和提升我国的AI竞争力；三是培育形成光电融合的完整产业链（设计、制造、装备、系统、应用等），占领信息产业发展高地。

（三）营造充满活力的光电融合创新生态

良好生态是创新的黑土地。光电融合涉及到芯片，资本和知识高度密集，生态系统复杂，有赖于政府和企业共同营造。一是借鉴国外发展微电子的成功经验，成立类似国际半导体技术路线图/国际器件与系统路线图的组织，以形成“产学研”各界对于关键技术识别、技术研究节奏、产业化时间点等的诸多共识，从而牵引创新链各环节、产业链各主体形成合力，确保“产学研用”各方有计划地高效协同。二是政府持续进行政策和资金支持，并引导资本和企业大力支持。光电融合需要深度创新，涉及到基础科学和制造技术的交叉，芯片、设备、系统等各层次架构变革，材料、结构、设计、工艺等各环节研发，各类软件工具（建模、仿真、设

计）和硬件装备，投入是创新活力的保障。三是吸取国外发展教训，注重提供公共创新平台设施。美国的先进微电子制造技术仅依赖于少数几家公司，绝大多数研究人员没有参与先进技术和工艺的研究机会和条件，间接导致其先进制程极为依赖进口。四是培养创新队伍，吸纳全球相关人才。

（四）培育和壮大有国际竞争力的光电融合创新主体

企业是产业主体，也是创新主体。我国缺少能影响行业发展的龙头企业，发展一批龙头企业，并形成覆盖全产业链的企业集群，对真正掌握信息产业发展主动权至关重要。需要构建国家战略牵引、资源要素集聚、良性生态支撑的三维体系，支持国内龙头企业发展壮大为有国际话语权和全球影响力的顶尖企业。一是要建立长周期的领军企业培育机制，支持龙头企业牵头组建能贯通国内科学、技术、产业链条的光电融合创新联合体，有效发挥领头雁在雁阵式集群中的关键作用。二是由政府和龙头企业共同设立基础研究联合基金，引导其在光电融合底层架构、根技术方面敢于投入，以科技创新跨越追赶者陷阱。三是支持龙头企业参与领域内国家战略规划制定，支持龙头企业主导或联合发起全球性的光电融合产业和技术联盟，由内而外逐步争取技术路径主导权、产业生态定义权。

利益冲突声明

本文作者在此声明不存在任何利益冲突或财务冲突。

Received date: November 14, 2025; **Revised date:** March 12, 2026

Corresponding author: Zhang Xinquan is a professor-level senior engineer from the State Key Laboratory of Optical Communication Technologies and Networks. His major research field is information optoelectronics. E-mail: zhangxq@cict.com

Funding project: Chinese Academy of Engineering project “Prospective and Reserve Advisory Research on Electronic and Information Engineering Technology in China” (2023-JB-14)

参考文献

- [1] 张新全, 余少华. 信息化前景和网络演进探讨 [J]. 光通信研究, 2021 (2): 1-6.
Zhang X Q, Yu S H. Discussion on the prospect of informatization and network evolution [J]. Study on Optical Communications, 2021 (2): 1-6.
- [2] Photonics' renaissance moment [EB/OL]. (2025-04-28)[2025-05-01]. <https://www.lightcounting.com/research-note/april-2025-ofc->

- 2025-photonics-renaissance-moment-398.
- [3] 余少华. 网络通信七个技术墙及后续趋势初探 [J]. 光通信研究, 2018 (5): 1.
Yu S H. The seven technical walls of network communications and the primary exploration of trends [J]. Study on Optical Communications, 2018 (5): 1.
- [4] 中国信息与电子工程科技发展策略研究中心. 中国电子信息工程科技发展研究: 信息光电子专题 [M]. 北京: 科学出版社, 2019.
China Research Center for Information and Electronic Engineering Science and Technology Development Strategy. Research on China's electronic information engineering science and technology development: Information optoelectronics [M]. Beijing: Science Press, 2019.
- [5] 张新全, 余少华. 光电融合破解带宽、能耗难题 [J]. 光通信研究, 2021 (5): 1–14, 23.
Zhang X Q, Yu S H. Address the challenges of bandwidth and power consumption through photonics-electronics convergence [J]. Study on Optical Communications, 2021 (5): 1–14, 23.
- [6] 张新全, 肖希, 余少华. 信息光电子“微电子化”技术进展和发展探讨 [J]. 光通信研究, 2023 (6): 1–10.
Zhang X Q, Xiao X, Yu S H. Technological progress of information optoelectronics microelectronization and discussion on its development [J]. Study on Optical Communications, 2023 (6): 1–10.
- [7] 中国信息与电子工程科技发展策略研究中心. 中国电子信息工程科技发展研究: 微电子光电子国内外发展态势研究 [M]. 北京: 科学出版社, 2024.
Research Center for Information and Electronic Engineering Science and Technology Development Strategy. Research on China's electronic information engineering science and technology development: Global and domestic development trends of microelectronics and optoelectronics [M]. Beijing: Science Press, 2024.
- [8] 谭旻. 光电融合芯片: 基本概念与未来展望 [J]. 光通信研究, 2025 (6): 32–44.
Tan M. Electronic-photonic heterogeneously-converging integrated circuits: Basic concepts and future prospects [J]. Study on Optical Communications, 2025 (6): 32–44.
- [9] Sakamoto T, Sato N, Segawa T. Photonics-electronics convergence technologies for disaggregated computing [J]. NTT Technical Review, 2021, 19(7): 58–64.
- [10] Shi W, Tian Y, Gervais A. Scaling capacity of fiber-optic transmission systems via silicon photonics [J]. Nanophotonics, 2018, 9(16): 4629–4663.
- [11] Orcutt J S, Gill D M, Proesel J, et al. Monolithic silicon photonics at 25 Gb/s [R]. Anaheim: Optical Fiber Communication Conference, 2016.
- [12] Silicon photonics reaches prime time [EB/OL]. (2019-04-26) [2025-05-22]. <https://www.eetimes.eu/silicon-photonics-reaches-prime-time/>.
- [13] Eppenberger M, Bonomi M, Moor D, et al. Compact optical TX and RX macros for computercom monolithically integrated in 45 nm CMOS [J]. Journal of Lightwave Technology, 2021, 39(21): 6869–6879.
- [14] Darbinian A, Ardalan S, Baehr-Jones T, et al. Monolithic Mach-Zehnder modulator bias controls in a 45 nm silicon photonics process [J]. IET Conference Proceedings, 2023, 2023(34): 1492–1495.
- [15] Sun C. Photonics for die-to-die interconnects: Links and optical I/O chiplets [R]. San Francisco: 2024 IEEE International Solid-State Circuits Conference (ISSCC), 2024.
- [16] Li A, Ma Q L, Xie Y J, et al. A 256 Gb/s electronic-photonically integrated transceiver in 45 nm CMOS [J]. Journal of Semiconductors, 2024, 45(7): 070501.
- [17] Atabaki A H, Moazeni S, Pavanello F, et al. Integrating photonics with silicon nanoelectronics for the next generation of systems on a chip [J]. Nature, 2018, 556(7701): 349–354.
- [18] Graef M. More than Moore white paper [R]. Santa Clara: 2021 IEEE International Roadmap for Devices and Systems Outbriefs, 2022.
- [19] 欧祥鹏, 杨在利, 唐波, 等. 2.5D/3D 硅基光电子集成技术及应用 [J]. 光通信研究, 2023 (1): 1–16.
Ou X P, Yang Z L, Tang B, et al. Silicon photonic 2.5D/3D integration technology and its applications [J]. Study on Optical Communications, 2023 (1): 1–16.
- [20] Moralis-Pegios M, Pitris S, Alexoudi T, et al. 4-channel 200 Gb/s WDM O-band silicon photonic transceiver sub-assembly [J]. Optics Express, 2020, 28(4): 5706–5714.
- [21] Xuan Z, Balamurugan G, Huang D N, et al. A 256 Gbps heterogeneously integrated silicon photonic microring-based DWDM receiver suitable for in-package optical I/O [R]. Kyoto: 2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits), 2023.
- [22] Levy C S, Xuan Z, Sharma J, et al. $8\text{-}\lambda \times 50$ Gbps/ λ heterogeneously integrated Si-ph DWDM transmitter [J]. IEEE Journal of Solid-State Circuits, 2024, 59(3): 690–701.
- [23] Daudlin S, Rizzo A, Abrams N C, et al. 3D-integrated multichip module transceiver for terabit-scale DWDM interconnects [R]. San Francisco: 2021 Optical Fiber Communications Conference and Exhibition (OFC), 2021.
- [24] Shekhar S, Bogaerts W, Chrostowski L, et al. Roadmapping the next generation of silicon photonics [J]. Nature Communications, 2024, 15(1): 751.
- [25] Reed G T, Mashanovich G, Gardes F Y, et al. Silicon optical modulators [J]. Nature Photonics, 2010, 4(8): 518–526.
- [26] Han C H, Jin M, Tao Y S, et al. Ultra-compact silicon modulator with 110 GHz bandwidth [R]. San Diego: 2022 Optical Fiber Communications Conference and Exhibition (OFC), 2022.
- [27] Xu M Y, Zhu Y T, Pittalà F, et al. Dual-polarization thin-film lithium niobate in-phase quadrature modulators for terabit-per-second transmission [J]. Optica, 2022, 9(1): 61.
- [28] Margalit N, Xiang C, Bowers S M, et al. Perspective on the future of silicon photonics and electronics [J]. Applied Physics Letters, 2021, 118(22): 220501.
- [29] Fatholouloumi S, Hui D, Jadhav S, et al. 1.6 Tbps silicon photonics integrated circuit and 800 Gbps photonic engine for switch co-packaging demonstration [J]. Journal of Lightwave Technology, 2021, 39(4): 1155–1161.
- [30] U.S. Naval Research Laboratory (NRL). 25 Technologies for the

- next 25 years (2023—2048) [R]. Washington DC: NRL, 2022.
- [31] 郝然. 对硅基光电子技术发展的思考 [J]. 中兴通讯技术, 2017, 23(5): 52–55.
- Hao R. Development of the silicon photonic technology [J]. ZTE Technology Journal, 2017, 23(5): 52–55.
- [32] Chi T, Li Z, Li N, et al. Laser integration on a photonic integrated circuit with high alignment accuracy for data transmission [R]. Orlando: 2023 IEEE 73rd Electronic Components and Technology Conference (ECTC), 2023.
- [33] Blum R. Integrated silicon photonics for high-volume data center applications [R]. San Francisco: Optical Interconnects XX, 2020.
- [34] Jones R, Doussiere P, Driscoll J B, et al. Heterogeneously integrated InP/silicon photonics: Fabricating fully functional transceivers [J]. IEEE Nanotechnology Magazine, 2019, 13(2): 17–26.
- [35] Diamantopoulos N P, Fujii T, Yamaoka S, et al. 16-channel directly modulated membrane III-V laser array on SiO₂/Si utilizing photon-photon resonance [J]. Journal of Lightwave Technology, 2024, 42(11): 3997–4005.
- [36] Shang C, Selvidge J, Hughes E, et al. A pathway to thin GaAs virtual substrate on On-Axis Si (001) with ultralow threading dislocation density [J]. Physica Status Solidi(a): Applications and Materials Science, 2021, 218(3): 202000402.
- [37] De Koninck Y, Caer C, Yulistira D, et al. GaAs nano-ridge laser diodes fully fabricated in a 300-mm CMOS pilot line [J]. Nature, 2025, 637(8044): 63–69.
- [38] Kawahara K, Tsuchizawa T, Yamamoto N, et al. 1.9-pJ/bit 0.5-mm² high-speed optical transmitter including silicon slow-light modulator and current-mode BiCMOS driver [R]. Frankfurt: ECOC 2024, 2024.
- [39] Chen S, Wang D, Yang M, et al. Enabling 3D heterogeneous integration of silicon photonics chiplet through UBM-Free Au-SnAg bonding for 2 Tb/s optical interconnects [R]. Shanghai: 2025 26th International Conference on Electronic Packaging Technology (ICEPT), 2025.
- [40] Nagashima K, Azuma T, Dong L, et al. An ultra-compact CPO transceiver based on a 1060-nm single-mode VCSEL array and multicore fibers [J]. Journal of Lightwave Technology, 2025, 43(13): 6176–6186.
- [41] Daudlin S, Lee S, Kilwani D, et al. Ultra-dense 3D integrated 5.3 Tb/s/mm² 80 micro-disk modulator transmitter [R]. San Diego: Optical Fiber Communication Conference (OFC) 2023, 2023.
- [42] Melek D T, Navinkumar R, Vandersand J, et al. A 0.29 pJ/b 5.27 Tb/s/mm UCIe advanced package link in 3 nm FinFET with 2.5D CoWoS packaging [R]. San Francisco: 2025 IEEE International Solid-State Circuits Conference (ISSCC), 2025.
- [43] Chua C, Zhou Y, Krishnamoorthy A V, et al. A 14.4 Tb/s silicon photonic I/O chiplet with 2.4 pJ/bit energy efficiency for UCIe-backed multi-chip integration [R]. San Francisco: IEEE International Solid-State Circuits Conference (ISSCC), 2024.
- [44] Sriram A, Shamsul A, Vincent C, et al. The Intel programmable and integrated unified memory architecture graph analytics processor [J]. IEEE Micro, 2023, 43(5): 78–87.
- [45] Sze V, Chen Y H, Yang T J, et al. Efficient processing of deep neural networks: A tutorial and survey [J]. Proceedings of the IEEE, 2017, 105(12): 2295–2329.
- [46] Wetzstein G, Ozcan A, Gigan S, et al. Inference in artificial intelligence with deep optics and photonics [J]. Nature, 2020, 588(7836): 39–47.
- [47] Kitayama K I, Notomi M, Naruse M, et al. Novel frontier of photonics for data processing—Photonic accelerator [J]. APL Photonics, 2019, 4(9): 090901.
- [48] Hua S, Divita E, Yu S, et al. An integrated large-scale photonic accelerator with ultralow latency [J]. Nature, 2025, 640: 361–367.
- [49] Xu Z H, Zhou T K, Ma M Z, et al. Large-scale photonic chiplet Taichi empowers 160-TOPS/W artificial general intelligence [J]. Science, 2024, 384(6692): 202–209.
- [50] Xue Z W, Zhou T K, Xu Z H, et al. Fully forward mode training for optical neural networks [J]. Nature, 2024, 632(8024): 280–286.
- [51] Zou P, Hu F, Zhao Y, et al. Optical computing accelerators: Principle, application, and perspective [J]. Frontiers of Physics, 2025, 20(3): 032302.
- [52] Chen Y K. Confluence of photonics and artificial intelligence [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2025, 31(3): 1–14.
- [53] Luo Y, Zhao Y F, Li J X, et al. Computational imaging without a computer: Seeing through random diffusers at the speed of light [J]. eLight, 2022, 2(1): 4.
- [54] Yildirim M, Dinc N U, Oguz I, et al. Nonlinear processing with linear optics [J]. Nature Photonics, 2024, 18(10): 1076–1082.
- [55] Fu T Z, Zang Y B, Huang Y Y, et al. Photonic machine learning with on-chip diffractive optics [J]. Nature Communications, 2023, 14: 70.
- [56] Shen Y C, Harris N C, Skirlo S, et al. Deep learning with coherent nanophotonic circuits [R]. San Juan: 2017 IEEE Photonics Society Summer Topical Meeting Series (SUM), 2017.
- [57] Choi C, Lee S Y, Mun S E, et al. Metasurface with nanostructured Ge₂Sb₂Te₅ as a platform for broadband-operating wavefront switch [J]. Advanced Optical Materials, 2019, 7(12): 1900171.
- [58] Komar A, Paniagua-Domínguez R, Miroshnichenko A, et al. Dynamic beam switching by liquid crystal tunable dielectric metasurfaces [J]. ACS Photonics, 2018, 5(5): 1742–1748.
- [59] Arbabi E, Arbabi A, Kamali S M, et al. MEMS-tunable dielectric metasurface lens [J]. Nature Communications, 2018, 9(2): 812.
- [60] Pai S, Sun Z H, Hughes T W, et al. Experimentally realized *in situ* backpropagation for deep learning in photonic neural networks [J]. Science, 2023, 380(6643): 398–404.
- [61] Feldmann J, Youngblood N, Karpov M, et al. Parallel convolutional processing using an integrated photonic tensor core [J]. Nature, 2021, 589(7840): 52–58.
- [62] Hua S Y, Divita E, Yu S S, et al. An integrated large-scale photonic accelerator with ultralow latency [J]. Nature, 2025, 640(8058): 361–367.
- [63] Xu X Y, Tan M X, Corcoran B, et al. 11 TOPS photonic convolutional accelerator for optical neural networks [J]. Nature, 2021, 589(7840): 44–51.
- [64] Wu N F, Sun Y X, Hu J T, et al. Intelligent nanophotonics: When machine learning sheds light [J]. eLight, 2025, 5(1): 5.
- [65] Chen Y T, Nazhamaiti M, Xu H, et al. All-analog photoelectronic chip for high-speed vision tasks [J]. Nature, 2023, 623(7985):

- 48–57.
- [66] Jouppi N, Kurian G, Li S, et al. TPU v4: An optically reconfigurable supercomputer for machine learning with hardware support for embeddings [R]. Orlando: The 50th Annual International Symposium on Computer Architecture, 2023.
- [67] Wang W Y, Ghobadi M, Shakeri K, et al. Rail-only: A low-cost high-performance network for training LLMs with trillion parameters [R]. Albuquerque: 2024 IEEE Symposium on High-Performance Interconnects (HOTI), 2024.
- [68] Cao W, Bu H M, Vinet M, et al. The future transistors [J]. *Nature*, 2023, 620(7974): 501–515.
- [69] 李尔平. 中国电子信息工程科技发展研究: 后摩尔时代微纳新器件电磁场专题 [M]. 北京: 科学出版社, 2025.
- Li E P. Electromagnetic fields in Micro/Nano new devices for the post-Moore era [M]. Beijing: Science Press, 2025.
- [70] Do Nascimento V C, Hwang S, Smith M J, et al. Multiphysics-informed ML-assisted chiplet floorplanning for heterogeneous integration [J]. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2025, 15(5): 959–973.
- [71] Xu D L, Ma Y C, Jin G F, et al. Intelligent photonics: A disruptive technology to shape the present and redefine the future [J]. *Engineering*, 2025, 46: 186–213.
- [72] Hernandez S, Peucheret C, Da Ros F, et al. End-to-end optimization of optical communication systems based on directly modulated lasers [J]. *Journal of Optical Communications and Networking*, 2024, 16(8): D29–D43.
- [73] Li J H, Li X, Wu L, et al. Ultra-compact scalable mode demultiplexers for high-speed optical interconnects *via* GPU-accelerated inverse design [J]. *Optics Express*, 2025, 33(21): 44908–44924.
- [74] Hossain M J, Reitano D, Rakin A S. Inverse design of silicon photonics components: A study from deep learning perspective [R]. Orlando: 2023 IEEE Photonics Conference (IPC), 2023.
- [75] Merchant A, Batzner S, Schoenholz S S, et al. Scaling deep learning for materials discovery [J]. *Nature*, 2023, 624(7990): 80–85.
- [76] The CHIPS Program Office vision for success: Two years later [EB/OL]. (2025-01-10)[2026-02-18]. <https://www.nist.gov/system/files/documents/2025/01/10/The%20CHIPS%20Program%20Office%20Vision%20for%20Success%20-%20Two%20Years%20Later.pdf>.
- [77] Xu D L, Ma Y C, Jin G F, et al. Intelligent photonics: A disruptive technology to shape the present and redefine the future [J]. *Engineering*, 2025, 46(3): 186–213.
- [78] 周林杰, 冉诗环, 原旗旗, 等. 硅光集成与光电融合: 后摩尔时代的关键使能技术 (特邀) [J]. *光学学报*, 2025 (17): 33–66.
- Zhou L J, Ran S H, Yuan Q Q, et al. Silicon photonic integration and photonics-electronics convergence: Key enabling technologies for the post-Moore era (invited) [J]. *Acta Optica Sinica*, 2025 (17): 33–66.
- [79] Shirao M, Lo P, Jiao Y. Evolution of photonics-electronics convergence technology—Can we make it, how and when? [R]. Sapporo: 2025 30th OptoElectronics and Communications Conference (OECC) and 2025 International Conference on Photonics in Switching and Computing (PSC), 2025.
- [80] Kanekiyo T, Shirai D, Inoue S. Release of NTT IOWN technology report 2024 [J]. *NTT Technical Review*, 2025, 23(2): 14–18.
- [81] IEEE. Heterogeneous integration roadmap 2024 edition [EB/OL]. (2024-11-10)[2026-03-10]. <https://eps.ieee.org/technology/heterogeneous-integration-roadmap/2024-edition/>.