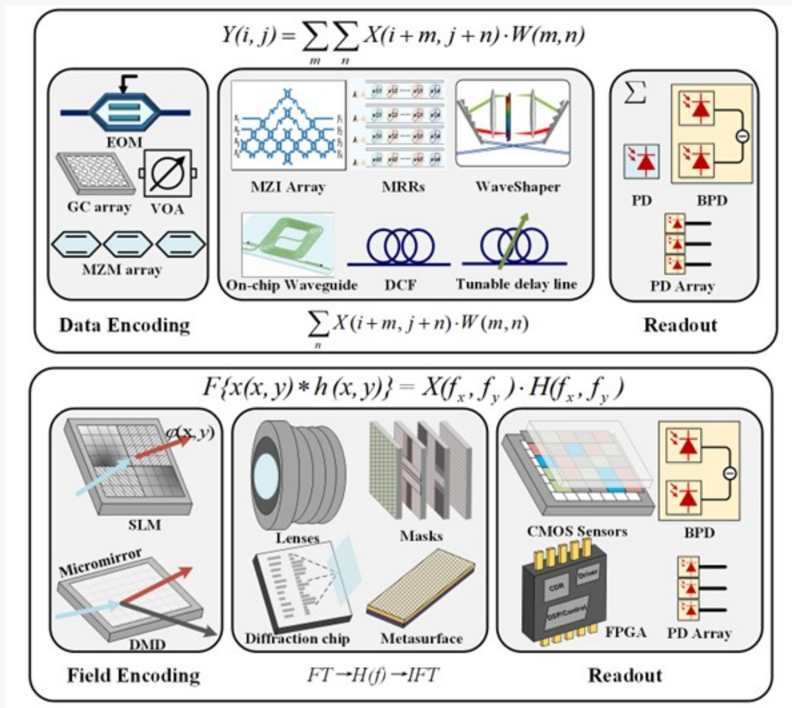
Optical Convolution Computation: Principle, Applications, Challenges

Review highlights advances, applications, and challenges in optical convolution for faster, more energy-efficient computing systems

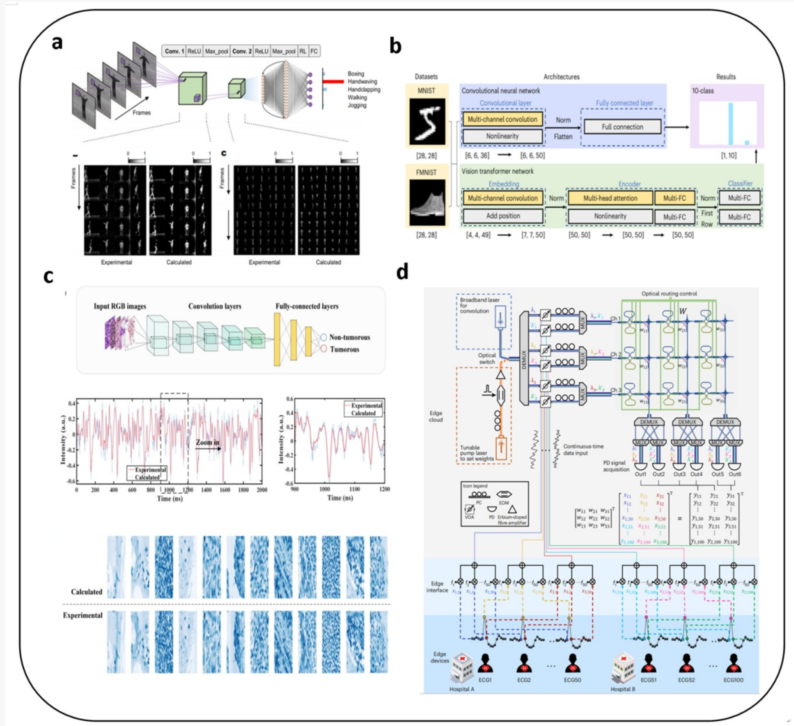


System-level overview of optical convolution accelerators

CHINA, August 28, 2026 /EINPresswire.com/ -- A review from Beijing University of Posts and Telecommunications organizes optical convolution computing into two approaches: definition-based and theorem-based methods. It highlights applications in signal processing, image analysis, and high-dimensional data processing, while identifying challenges such as energy efficiency, scalability, stability, and complete Convolutional Neural Network support. The review also emphasizes hybrid opto-electronic systems as a practical path toward deployable optical computing hardware.

Section 1: Background

Convolution is the core computational operator in Convolutional Neural Networks (CNNs), typically accounting for over 80% of the total inference workload. As CNN backbones continue to scale in depth and input resolution, the massively parallel multiply–accumulate (MAC) operations within convolutional layers—involving extensive fan-in, fan-out, and highly regular sliding-window dataflow—have become the dominant bottleneck for end-to-end throughput and energy efficiency. Unlike conventional electronic processors that are constrained by serialized execution and finite interconnect bandwidth, optical computing can structurally match the linear-transform nature of convolution at the physical level: the parallel propagation of light enables large-scale signal broadcast and fan-out; wavelength, spatial-mode, polarization, and other multiplexing dimensions provide native parallelism; and optical interference and field superposition directly implement weighted summation upon photodetection. This algorithm-hardware structural correspondence—rather than raw device speed alone—constitutes the primary motivation driving optical convolution accelerator research.



Application of optical convolutional computing in synchronous processing of high-dimensional signals and high-throughput medical image diagnosis.

In recent years, the rapid maturation of integrated photonic platforms—including silicon photonics, microcomb sources, high-speed modulators, and Mach–Zehnder interferometer (MZI) meshes—has enabled optical convolution to transition from bulky laboratory prototypes toward scalable, chip-level systems. However, the field remains fragmented: physical implementations span a heterogeneous range from free-space 4f Fourier systems to on-chip microring weight banks, CNN operator support is often incomplete (lacking full stride, padding, and multi-channel semantics), and evaluation metrics are inconsistent across studies, making fair comparisons difficult. To address this gap, Prof. Kun Xu's team at the State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, (BUPT), China, systematically reviewed the field. Grounded in two fundamental mathematical principles—the convolution definition and the convolution theorem—the review systematically organizes the field into a unified taxonomy, clarifies shared system-level signal chains in hybrid opto-electronic implementations, and provides comparable performance benchmarks to guide future development.

Section 2: Summary of Review Content

The research team led by Prof. Xu at BUPT was invited to publish a comprehensive review article entitled "Optical Convolution Computation: Principle, Applications, Challenges" in Volume 2 of the journal Intelligent Opto-Electronics on June 29, 2026. The review systematically organizes the diverse landscape of optical convolution computing into two major paradigms rooted in distinct mathematical principles.

The first paradigm, definition-based optical convolution (definition-based), directly implements the sliding-window MAC operations of CNNs in the physical domain. In this approach, optical modulation—via amplitude, phase, or polarization control—performs real-valued weighting, while optical field superposition upon photodetection realizes accumulation, creating a direct physical isomorphism to the convolution operator. Representative architectures include: dimension-interleaved schemes that exploit time, wavelength, space, and mode degrees of freedom for massive parallelism; microring resonator weight banks using wavelength-division-multiplexed broadcast-and-weight; coherent MZI meshes that map matrix multiplication onto programmable phase networks via singular value decomposition; and spatial-projection architectures that generate shifted replicas of the input field through gratings or beam splitters to realize sliding-window convolution in a single optical exposure.

The second paradigm, theorem-based optical convolution (theorem-based), leverages the convolution theorem to transform spatial-domain convolution into pointwise multiplication in the frequency domain, exploiting the intrinsic Fourier-transform capability of lenses and engineered photonic structures. The review covers classical 4f Fourier-optics processors, where convolution kernels are encoded as phase or amplitude masks at the Fourier plane; on-chip nanophotonic convolvers using inverse-designed achromatic metalenses for compact Fourier transforms; ultracompact meta-imagers that replace bulky 4f relays with programmable metasurface stacks; and frequency-comb-based spectral convolution schemes that perform multiply-accumulate through coherent detection of modulated comb lines, with one silicon-photonic implementation experimentally reaching the TOPS regime in a single computing cell.

Beyond architecture-level categorization, the review organizes applications by data dimensionality: one-dimensional radio frequency and radar signal processing exploiting the ultra-high temporal bandwidth of photonics for nanosecond-latency sensing; two-dimensional image classification, edge detection, and medical image reconstruction, where multiple demonstrations have achieved accuracy comparable to that of electronic implementations; and three-dimensional to higher-order tensor processing—including video action recognition, multi-channel ECG analysis, and biomedical diagnostics—where optical convolution offers the unique advantage of native high-dimensional processing without costly data reshaping. The review underscores that the field is now progressing from isolated operator demonstrations toward CNN-complete accelerators, and identifies the core challenges ahead: establishing standardized full-chain energy-efficiency metrics that encompass DAC/ADC interfaces, modulator drivers, and calibration overhead; achieving scalable integration with robust compensation for phase drift and thermal crosstalk; and advancing algorithm-hardware co-design. Critically, the review emphasizes that a hybrid opto-electronic paradigm—in which photonics executes bandwidth-intensive linear operators while electronics handles data marshaling, memory access, nonlinear activation, and closed-loop control—represents the most practical and viable path toward deployable optical convolution hardware.

Section 3: Outlooks

Optical convolution computing stands at a pivotal juncture, transitioning from principle validation toward deployable, system-ready accelerators. The review identifies the critical challenges ahead: achieving CNN-complete operator support that goes beyond single-layer demonstrations to encompass stride, padding, multi-channel processing, and cascaded layers; establishing full-chain energy-efficiency standards that transparently account for DAC/ADC interfaces, modulator drivers, and calibration overhead; and resolving scalability and stability bottlenecks arising from phase drift, thermal crosstalk, and fabrication variability in dense photonic circuits. Looking forward, optical convolution is poised to deliver early and compelling advantages in bandwidth-critical and latency-sensitive domains such as real-time edge intelligence and ultra-high-throughput signal processing.

Section 4: Introduction to the Research Group

This review was led by Prof. Kun Xu at the State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications (BUPT), China. Prof. Xu's team specializes in integrated photonic computing, optical neural networks, microcomb-based signal processing, and high-speed optical communication systems. Core team members Prof. Xu and Associate Prof. Yunping Bai have long been dedicated to integrated photonic computing, optical neural networks, microcomb-based signal processing, and high-speed optical communication systems. Their recent contributions include TOPS-speed photonic convolution accelerators, microcomb-driven photonic processing units, complex-valued optical convolutional accelerators, and photonic tensor cores for high-dimensional data processing.

Prof. David J. Moss from the Optical Sciences Centre, Swinburne University of Technology, Australia, is a collaborator on this review, with deep research expertise in integrated photonics and nonlinear optics.

The group has published extensively in high-impact journals and international conferences including Nature, Nature Photonics, Nature Communications, Light: Science & Applications, and Laser & Photonics Reviews. Their work has advanced optical computing from proof-of-concept demonstrations toward systematic, deployable architectures, with a particular focus on hybrid opto-electronic paradigms. Prof. Xu's team maintains active international collaborations and is equipped with state-of-the-art silicon photonic experimental platforms.

***

Reference

Title of original paper: Optical convolution computation: principle, applications, challenges

Journal: Intelligent Opto-Electronics
DOI: https://doi.org/10.67704/ioe.2026.260003

Funding information
This work was financially supported by the National Key Research and Development Program of China (No.2024YFB2808300); the Scientific Research Innovation Capability Support Project for Young Faculty (No. ZYGXQNJSKYCXNLZCXM-15); National Natural Science Foundation of China (NSFC) (No,62301074; 62135009).

Siyi Ma
Institute of Optics and Electronics
+86 191 3618 6296
oej_publishing@ioe.ac.cn
