Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

In-molecule machine learning with microfluidic integrated DNA neural computing

Chunhai Fan

fanchunhai@sjtu.edu.cn

Shanghai Jiao Tong University https://orcid.org/0000-0002-7171-7338

Mengyao Cao

East China Normal University

Xiewei Xiong

East China Normal University

Jianing Wang

East China Normal University

Na Li

East China Normal University

Bowen Li

East China Normal University https://orcid.org/0009-0004-7863-8247

Yun Zhu

East China Normal University

Tong Zhu

East China Normal University https://orcid.org/0000-0001-7472-3736

Fei Wang

Shanghai Jiao Tong University

Longqian Xu

Suzhou Institute of Biomedical Engineering Technology, Chinese Academy of Sciences

Jiajian Ji

Suzhou Institute of Biomedical Engineering Technology

Qunyan Yao

Fudan University https://orcid.org/0000-0003-3586-1911

Li Li

East China Normal University https://orcid.org/0000-0001-8494-4997

Hao Pei

East China Normal University https://orcid.org/0000-0002-6885-6708

Chunyu Chang

University of Electronic Science and Technology of China

Article

Keywords:

Posted Date: October 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-5440528/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Microfluidic integrated DNA neural computing

Mengyao Cao[1], Na Li[1], Xiewei Xiong[1]*, Sichen Xing[1], Jianing Wang[1], Bowen Li[1], Yun Zhu[1], Tong Zhu[1], Fei Wang[2], Siyi Hu[4], Hanbin Ma[5], Longqian Xu[5], Chunyu Chang[6], Jiajian Ji[4], Qunyan Yao[7], Li Li[1], Hao Pei[1]*, Chunhai Fan[2,3]*

[1]Shanghai Key Laboratory of Green Chemistry and Chemical Processes; Shanghai Center of Brain-inspired Intelligent Materials and Devices; Shanghai Frontiers Science Center of Molecule Intelligent Syntheses; School of Chemistry and Molecular Engineering, East China Normal University; 500 Dongchuan Road, Shanghai 200241, China

[2] State Key Laboratory of Synergistic Chem-Bio Synthesis, School of Chemistry and Chemical Engineering, New Cornerstone Science Laboratory, Frontiers Science Center for Transformative Molecules, Zhang Jiang Institute for Advanced Study and National Center for Translational Medicine, Shanghai Jiao Tong University, Shanghai 200240, China

[3]Renji Branch of National Center for Translational Medicine, Institute of Molecular Medicine, Shanghai Key Laboratory for Nucleic Acids Chemistry and Nanomedicine, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China

[4]Suzhou Institute of Biomedical Engineering Technology, Chinese Academy of Sciences,88 Keling Road, Suzhou, 215163, China

[5]School of Medicine, University of Electronic Science and Technology of China, Chengdu, 611731, China

[6]College of Computer Science and Technology, Jilin University, Changchun, 130012, China

[7]Department of Gastroenterology and Hepatology, Zhongshan Hospital, Fudan University, Shanghai 200032, P.R. China

*Corresponding authors. Email: xwxiong@chem.ecnu.edu.cn; peihao@chem.ecnu.edu.cn; fanchunhai@sjtu.edu.cn;

Abstract

Molecular computing systems hold unique potential for processing biological information, but their development has been fundamentally limited by the physical bottlenecks of homogeneous reaction systems in that molecular crosstalk and uncontrollable reaction dynamics severely constrain computational scale and complexity. Here, we report a large-scale DNA-encoded microfluidic integrated circuit (DMIC) that encapsulates DNA computational elements into droplet-based molecular chiplets and manipulates them via a high-density programmable electrode array. This hybrid bioelectronic platform establishes a fluidic von Neumann architecture with memory, processor, controller, and I/O units, which compiles algorithms into physical droplet operations, thereby transforming molecular computation from a diffusion-limited continuum into a reconfigurable, non-equilibrium integrated fluidic circuit. Using this approach, we demonstrate a 144-neuron convolutional neural network assembled from over 2,000 molecular chiplets, classifying 36 categories of handwritten symbols with >80% accuracy. The DMIC platform also executes on-chip supervised learning

and performs multi-class cancer diagnosis from clinical serum samples. Our work provides a generalizable engineering framework for scalable molecular machine learning and opens pathways toward intelligent, autonomous bio-integrated systems.

Introduction

Machine learning algorithms are highly effective for recognizing patterns within high-dimensional biological data[1,2]. *In-silico* machine learning, leveraging the computational power of silicon-based hardware, can analyze massive multi-omics data to uncover disease-associated molecular features, thereby facilitating precision medicine[3-5]. Similarly, *in-molecular* machine learning, which emulates machine learning functions through chemical reaction networks, can identify molecular signatures to derive insights into disease states[6-9]. By processing biological information directly within its native molecular context[10-14], this approach eliminates the need for extensive data conversion, facilitating more efficient and real-time analysis, and showing significant potential for intelligent diagnostics and point-of-care testing. For example, DNA-based chemical reactions are being explored to develop biocompatible computing systems to perform machine learning tasks, such as mimicking artificial neurons[15-21]. These intelligent molecular systems have also been demonstrated as molecular multi-gene classifiers for diagnosing cancers and respiratory infections[22-26]. However, scaling such molecular systems faces a fundamental physical constraint in homogeneous solution. As network complexity grows, the escalating interplay of unintended molecular interactions (crosstalk) and stochastic reaction dynamics severely limits both the scale and reliability of computation.

DNA neural computing has been primarily advanced through molecular engineering approaches in homogeneous-phase reaction systems[6,8,27], such as DNA gate design and sequence optimization[16,17,28-32], generating significant interest in addressing challenges like molecular signal propagation and nonlinear dynamics. These efforts aim to construct reliable DNA-based neurons and organize them into interconnected

networks, enabling the implementation of diverse neural network algorithms[15-19,21,33]. For example, there have been experimental demonstrations of DNA-based reaction networks, including Hopfield associative memory[15], winner-take-all circuits[16,20,21], convolutional neural networks[17], and deep artificial neural networks powered by enzymes[18]. Furthermore, the persistent inherent stochasticity and complex dynamics of molecular reactions[15,28,34-36] remain a critical bottleneck for constructing synthetic molecular systems capable of meeting demanding computational requirements. These difficulties are particularly pronounced when building large-scale neural networks with complex topologies. The inability to ensure robust control over massively parallel biochemical reactions limits both the scalability and computational capacity of these systems. Therefore, addressing these severe limitations necessitates a fundamental paradigm shift in the strategies of molecular neural computing to fully unlock its computational potential.

Biological systems have evolved to manage complex biochemical reaction networks that support their computational abilities through compartmentalization and integration into fluidic networks with dynamic connections[37-40]. This is evident in how biological organisms have cells that are self-organized for isolating highly reactive chemicals and dynamically reconfiguring network connectivity for distinct tasks. Inspired by the principles of modern electronics and biological compartmentalization, we introduce a hybrid bioelectronic platform that establishes an integrated fluidic circuit for molecular computation. This DNA-encoded microfluidic integrated circuit (DMIC) encapsulates DNA computational elements into discrete, droplet-based molecular chiplets, a concept inspired by chiplets in in modern semiconductor engineering, where massive monolithic systems are disaggregated into standardized, modular blocks to definitively overcome physical scaling limits. These chiplets are manipulated within a microfluidic device [41,42] by a high-density, programmable electrode array, which functions as a dynamic interconnection network. This setup effectively instantiates a fluidic von Neumann architecture, complete with memory (chiplets in reservoirs), processor (reaction chambers), controller (external electronics), and I/O units. In this architecture, a computational

algorithm is compiled into a sequence of physical operations, retrieving, transporting, and merging specific chiplets to form a purpose-built network. This approach spatially isolates parallel reactions to eliminate crosstalk and replaces diffuse molecular encounters with deterministic, electronically orchestrated assembly. It thereby transforms the computational substrate from a diffusion-limited continuum into a structured, non-equilibrium, and programmable information-processing system, clearing the fundamental physical barriers for scalable molecular computing.

Architecture of the DNA-encoded microfluidic integrated circuit (DMIC)

Traditional molecular computing systems, typically executed in "one-pot" homogeneous reaction environments, are fundamentally limited by physical bottlenecks such as uncontrolled molecular crosstalk and stochastic, diffusion-limited reaction dynamics. As the complexity of chemical reaction networks scales, these inherent constraints severely compromise computational reliability and throughput. To overcome these challenges, we introduce a DNA-encoded microfluidic integrated circuit (DMIC) that shifts the computing paradigm from a diffusion-limited continuum into a programmable, deterministic information-processing system (Fig. 1a). The core of the DMIC platform is the encapsulation of DNA computational elements into discrete, droplet-based molecular chiplets. By isolating parallel biochemical reactions within these chiplets, we eliminate off-target crosstalk and replace random molecular encounters with electronically orchestrated assemblies. The hardware establishes a fluidic von Neumann architecture, consisting of four functionally distinct units: memory (a reservoir storing the chiplet library), processor (reaction chambers for reaction execution), controller (external electronics driving the electrode array), and I/O (interfaced for algorithm input and signal readout) (Fig. 1b). In this architecture, a user-defined algorithm is compiled into a high-level instruction set, which is subsequently translated into low-level electronic signals to drive automated droplet manipulation and seamless integration of computing elements into complex, adaptive DNA-based networks

(Fig. 1c). The system utilizes a high-density programmable electrode array, featuring up to 720 × 1280 pixels in our large-scale implementation, to transport, merge, and mix droplets with nanoliter precision (e.g., 2.20 nL per chiplet). This integrated fluidic circuit approach spatially isolates parallel reactions to eliminate crosstalk while enabling scalable molecular machine learning through the reconfigurable assembly of standardized modular components.

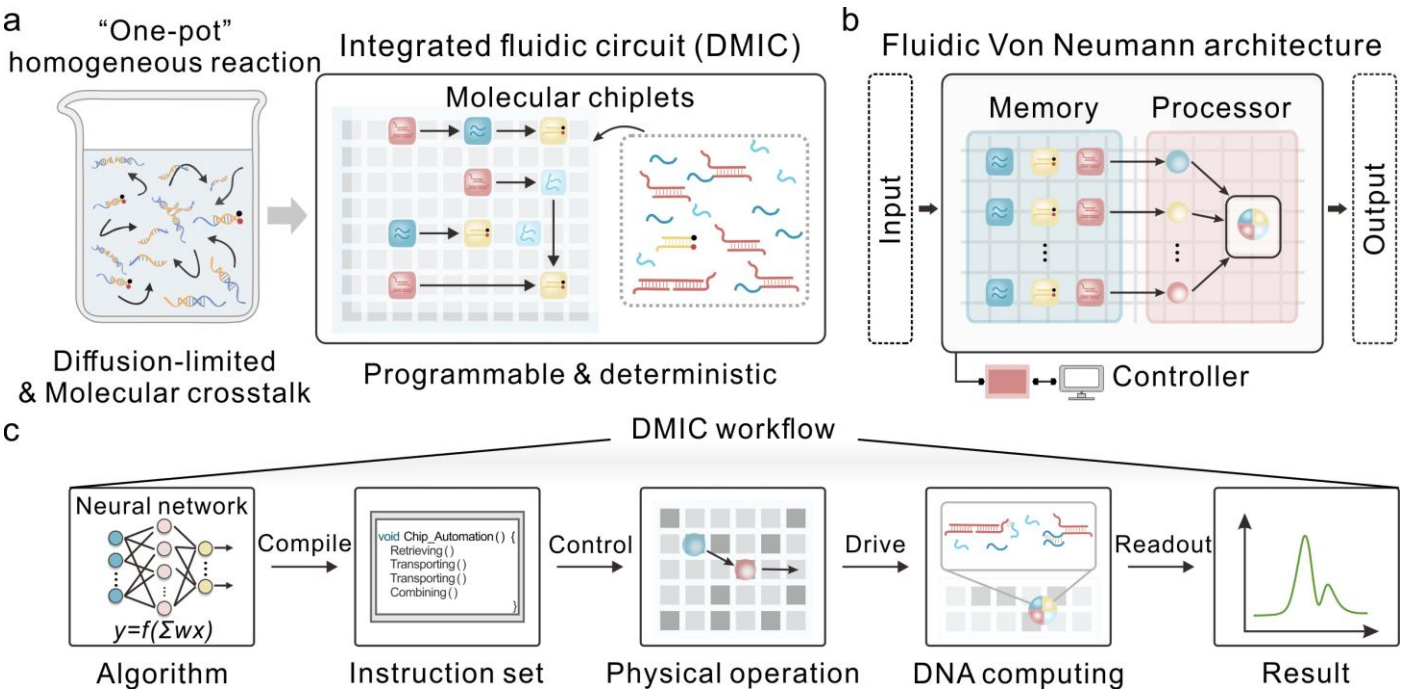


Fig. 1| Design and programmable workflow of the DMIC platform a, Comparison between traditional and integrated molecular computing. Left: Conventional one-pot systems rely on diffusion-limited, stochastic molecular interactions. Right: The DMIC platform utilizes an integrated fluidic circuit to encapsulate DNA computational elements into molecular chiplets, enabling a programmable, order-deterministic reaction environment. **b,** Fluidic von Neumann architecture. The system is organized into a memory for chiplet storage, a processor for reaction and computation, and an external controller (PC and driver electronics) that orchestrates the overall operation. **c,** DMIC operational workflow. The automated process comprises five stages: (1) User-defined algorithms are established; (2) Algorithms are compiled into an instruction set; (3) Electronic signals drive droplet manipulation on the electrode array; (4) DNA computation is executed on-chip through physical chiplet assembly; (5) Computation results are acquired via fluorescence readout.

High-precision droplet manipulation and the instantiation of molecular intelligence

The transition from a theoretical fluidic von Neumann architecture to a functional integrated computing system relies on the high-precision manipulation of matter within the DMIC hardware. Our platform utilizes

a programmable active electrode array to coordinate the movement of molecular chiplets, droplet-encapsulated DNA computational elements, across a physically structured environment (Fig. 2a). We developed a comprehensive library of molecular chiplets, each consisting of a droplet compartmentalizing distinct DNA computational elements: DNA neurons, inputs, weight modulators, and reporters (Fig. S1). By dynamically routing the chiplets across an integrated electrode grid, the system creates rewireable molecular pathways that interconnect molecular components, analogous to wiring in electronic circuits. The hardware used for initial experimental validation incorporates 138 electrodes, 12 reservoirs, and 2 reaction chambers (Fig. S2), facilitating the spatial isolation of distinct chemical species while ensuring their deterministic assembly on demand. By translating single algorithmic steps into specific instruction sets, we implemented a sequence of elementary physical operations including aliquoting, transporting, splitting, and merging & mixing (Fig. 2b and Figs. S3-5). These operations serve as the physical implementation of computational instructions, allowing the system to retrieve molecular chiplets with microliter precision (Fig. S4). For instance, DNA computational elements Inputs I_i are stored as standardized chiplets *e*Drop(I_i) in the reservoir and subsequently routed to reaction chambers to instantiate specific network topologies. This integrated fluidic circuit replaces the stochastic diffusion of bulk-solution systems with electronically orchestrated events, ensuring that each computational step is executed with high fidelity. The reliability of this programmable approach was validated first through the implementation of a single DNA neuron (Fig. 2c and Figs. S6-11). Using nine molecular chiplets, the system performed weighted-sum calculations and non-linear activation by merging specific droplets into a 2.75 μL reaction mixture (Fig. 2d). Characterization across 36 weight configurations showed that the on-chip outputs aligned closely with theoretical models, confirming that the physical operations successfully emulate the intended neuronal functions (Fig. 2e). This precise control over individual units provides the necessary technical foundation for scalable molecular computation, enabling the assembly of these building blocks into complex, multi-layered networks.

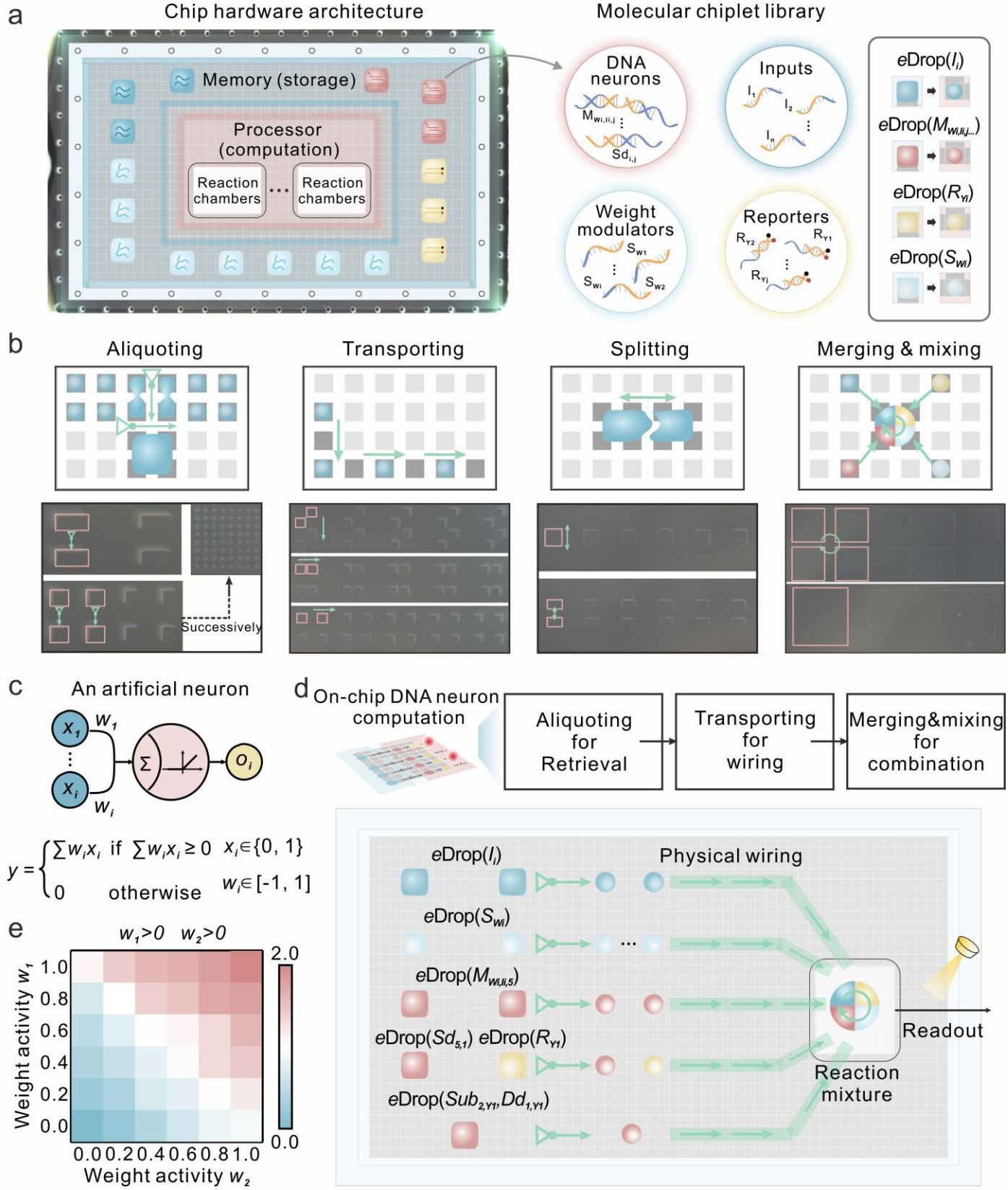


Fig. 2| Core hardware architecture and operations of the programmable DMIC. a, Hardware foundation of the integrated fluidic circuit. The programmable active electrode array orchestrates the spatiotemporal movement of molecular chiplets across a physically structured environment. **b,** The four elementary physical operations (aliquoting, transporting, splitting, merging & mixing). These deterministic, electronically orchestrated movements serve as the physical implementation of computational instructions. **c-e,** Instantiation of a single DNA neuron computation. By translating algorithmic steps into specific instruction sets, individual molecular chiplets are routed and merged to execute weighted-sum calculations and non-linear activation, establishing the technical foundation for scalable molecular computation. The x- and y- axis represent the

value of weights (w = 0.0, 0.2, 0.4, 0.6, 0.8 or 1) implemented using the weight modulator chiplets, with colours indicating the corresponding output signal intensity. The color scale (0–2) represents the normalized analog fluorescence outputs mapped to the corresponding theoretical output values for the tested weight combinations. The upper limit of 2 is determined by the maximum theoretical output values accessible in this experiment. The standard concentration is 50 nM (1×).

System-level integration of multi-DNA neuron networks on the DMIC platform

To explore the scalable potential of the integrated fluidic circuit, we developed an enlarged DMIC chip designed for the system-level integration of multiple DNA neurons (Fig. 3a and Fig. S13). This advanced hardware incorporates 305 electrodes, 28 reservoirs, and 4 reaction chambers, providing the necessary infrastructure for constructing sophisticated on-chip neural networks through automated programmable routing. By encapsulating distinct DNA neurons into discrete molecular chiplets, we established a modular toolkit that allows DNA neurons to be immobilized in reservoirs and dynamically interconnected via the electrode array (Fig. 3b). This architecture effectively decouples network topology from the physical hardware, as connections and weights can be redefined simply by reprogramming the droplet transport pathways. A significant advantage of this programmable approach is that droplet compartmentalization allows for the deployment of identical DNA sequences as distinct functional units in separate chambers (Fig. S14). This bypasses traditional sequence design constraints and maximizes parallelism, enabling the assembly of large-scale networks with high resource efficiency.

Experimental results demonstrated the integration of up to 26 DNA neurons into multilayered networks (Fig. 3c-h and Figs. S15-19). The system executed 13 distinct logical functions, ranging from simple logic cascades to complex algorithmic computation. The platform's flexibility was further evidenced by the implementation of advanced algorithmic circuits. For example, a ten-DNA neuron network was programmed to function as a square-root circuit (Fig. 3g and Fig. S18), while a 26-DNA neuron network performed fixed-point square-root calculations (Fig. 3h and Fig. S19). Most notably, we achieved a high-density 26-neuron network composed of four parallel subnets while requiring only eight unique DNA neurons (Fig. S20), demonstrating a significant leap in molecular resource optimization. Furthermore, our platform supported

concurrent processing, with multiple networks operating synchronously or asynchronously (Fig. S21), thereby enabling more complex algorithmic integration. As the network capacity scales exponentially, with n DNA neurons supporting 2^n logical functions, this fluidic von Neumann architecture represents a major shift toward hardware-reprogrammable and scalable molecular computation.

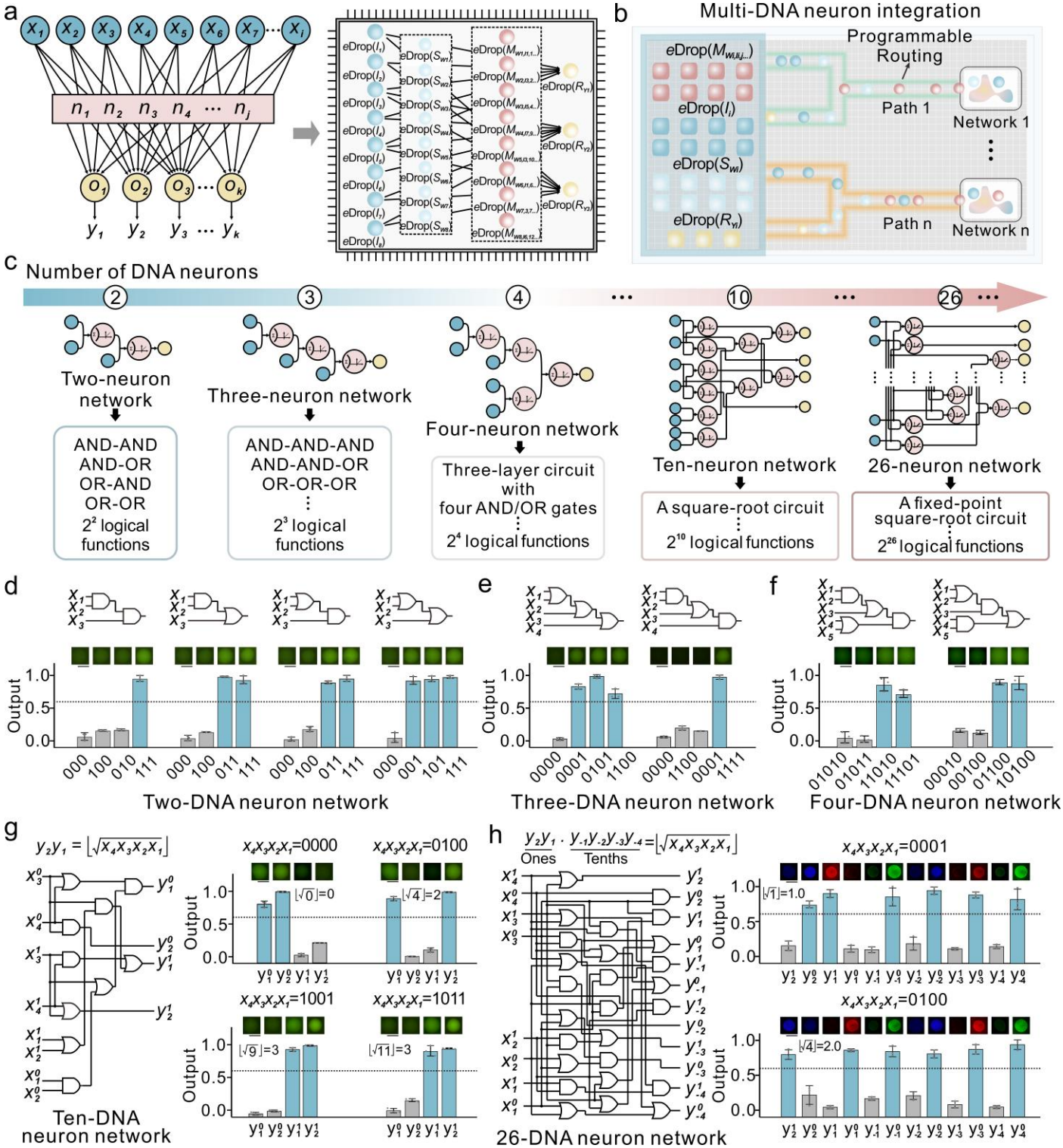


Fig. 3| System-level integration and reconfigurable connectivity of multi-DNA neuron networks on DMIC. a, Schematic of a general multi-layer neural network architecture with programmable connections. **b,** Deployment of molecular chiplets as standardized building blocks on the integrated fluidic circuit. This modular architecture completely decouples network topology from the physical hardware; inter-neuronal connections are dynamically established via electronically routed droplet pathways from designated reservoirs to target reaction chambers. **c,** Representation of diverse connectivity patterns. By merely reprogramming the transport pathways within the fluidic von Neumann architecture, the platform dynamically configures network topologies ranging from two-DNA neuron to 26-DNA neuron configurations. **d–h,** Experimental characterization of multi-DNA neuron network topologies. The system effortlessly transitions from simple logic cascades to complex algorithmic circuits: a two-layer logic cascades (**d**), the three-layer logic cascades (**e, f**), a square-root circuit utilizing the dual-rail convention (**g**), and a fixed-point square-root circuit utilizing the dual-rail convention (**h**). Experimental details: Fluorescence levels were recorded at 6 h (**d**) or 12 h (**e–h**) after input addition. Note that the require *e*Drop(I_i) were first merged and then routed to the target reaction chamber to minimize timing variability in input delivery. Input strands were introduced at 0× (logic 'OFF') or 1× (logic 'ON'), and an output is defined as logic 'ON' when its normalized signal exceeds the 0.50 threshold (dotted line). For the 26-neuron network (**h**), three distinct fluorophores (FAM, Atto425, and Texas Red) were utilized to characterize three independent outputs within a single reaction chamber.

Large-scale DNA-based convolutional neural network for 36-category pattern recognition

To demonstrate the scalable complexity of the integrated fluidic circuit, we developed a large-scale DMIC chip capable of the large-scale integration of hundreds of DNA neurons (Fig. 4a). This advanced hardware incorporates a 720 × 1280 active electrode array (Fig. S22), enabling the simultaneous configuration of hundreds of molecular chiplets with significantly higher integration density than previous iterations. The expanded capacity of this programmable platform supports the implementation of a convolutional neural network (ConvNet) designed to recognize 36 distinct handwritten symbols, including 10 digits and 26 letters from the EMNIST dataset (Fig. S23). The network architecture (Fig. 4b) processes 20 × 20 binary inputs through a convolutional layer with 36 kernels (10 × 10), followed by ReLU activation and global sum pooling, ultimately yielding 36 prediction outputs. The network comprises 144 sparsely connected neurons, 3,600 weight parameters, and 14,400 physical connections, achieving 80.64% accuracy on test data (Fig. S24).

For on-chip execution (Fig. 4c and Fig. S25), signed kernel weights were mapped to weight modulator chiplets (*e*Drop(S_{Wi})), with specific weight values represented by droplet counts. Weight sharing was achieved by applying identical *e*Drop(S_{Wi}) configurations across all receptive regions assigned to the same kernel (Fig. S26a). Binary input pixels were encoded by the presence or absence of corresponding *e*Drop(I_i) (Fig. S26b).

DNA neurons, each responsive to a 100-bit receptive region of input pattern, were encapsulated into discrete molecular chiplets for on-demand retrieval and deployment (Figs. S27 and 28). The pooling operation was implemented using conversion gates (eDrop(T_i)) to facilitate signal integration and dimensionality reduction (Fig. S29) To execute large-scale ConvNet computations on DMIC, which require the immense coordination of 36 parallel droplet networks organized from 144 DNA neurons, we utilized the Safe Interval Path Planning (SIPP) algorithm to optimize routing and avoid path conflicts (Fig. S30). This allowed for the simultaneous manipulation of approximately 2,304 droplets, transforming the fluidic von Neumann architecture into a high-throughput information-processing system. Experimental results (Fig. 4d) validated the DNA-based ConvNet's functionality, with a comprehensive confusion matrix confirming correct pattern classification across all 36 categories, demonstrating that the integrated fluidic circuit can reliably execute high-dimensional machine learning tasks. This achievement represents a qualitative shift in molecular computing, proving that the programmable and scalable nature of DMIC hardware is the key to overcoming the complexity bottlenecks of traditional DNA-based reaction systems.

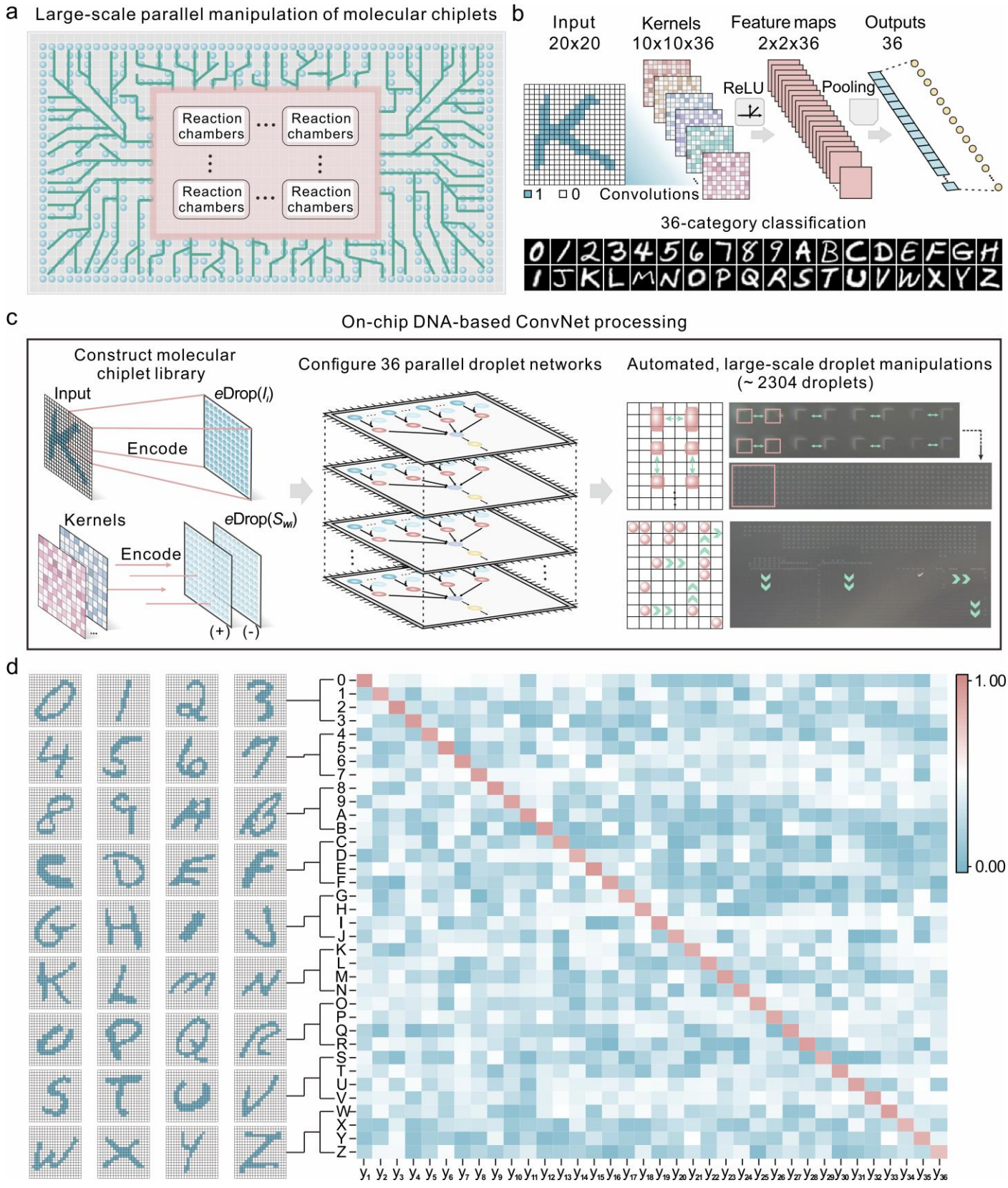


Fig. 4| Large-scale on-chip DNA-based ConvNet for 36-category handwritten symbol recognition. a, Architecture and routing map of the large-scale integrated fluidic circuit. Featuring a 720 × 1280 active-matrix electrode array, the platform utilizes the SIPP algorithm to manage the simultaneous, collision-free routing of thousands of molecular chiplets (~2,304 droplets). **b,** The 144-neuron ConvNet architecture and target symbols. Top: The network architecture designed to process 20 × 20 binary inputs through convolutional and pooling layers. Bottom: The 36 target handwritten symbols from the EMNIST database used to validate the system's pattern recognition capabilities. **c,** Implementation of the DNA-based ConvNet utilizing molecular chiplets with weight-sharing strategies. **d,** Experimental classification of 36 example handwritten symbols. The

color scale (0–1) represents the normalized experimental classification outputs, with the maximum response in each dataset set to 1. A threshold of 0.5 was applied to the normalized outputs at the final readout stage to define ON/OFF states, with each input pattern assigned to the class corresponding to the unique ON output.

Adaptive intelligence and clinical diagnostic applications of the DMIC platform

The DMIC platform further demonstrates its capacity for adaptive intelligence by implementing supervised learning directly within the physical molecular environment. Unlike purely *in silico* approaches, this system incorporates physical molecular networks into a closed-loop training process using a backpropagation rule[43] (Fig. 5a). We designed a convolutional neural network (ConvNet) with a 4×3 input layer, a 4×1 kernel, a ReLU activation layer, and a 3-class output to classify 12-bit binary patterns representing handwritten symbols 'L', 'T', and 'J' (Figs. S31-33). This programmable learning system utilized a hardware configuration of 227 electrodes, 19 reservoirs and 9 reaction chambers to execute forward propagation on-chip (Fig. S34). Following off-chip gradient calculations, the platform automatically reconfigured the routing of weight modulator chiplets (eDrop(S_{Wi})) to instantiate updated weights. Experimental results confirmed successful on-chip training across all cases, with 100% recognition accuracy across 39 training patterns within nine epochs (Figs. 5b–d). Across a range of initial kernel-weight configurations and learning rates (1.0 to 0.5), kernel weights evolved from random to optimized states—progressively improving feature discrimination—and the system achieved 100% accuracy under all tested conditions, with faster convergence at higher learning rates (Figs. S35-42). Classification results matched those of digital simulations, confirming operational robustness (Fig. S43). These results demonstrate the platform's ability to self-optimize through the iterative assembly of integrated fluidic circuits.

Expanding toward high-impact clinical utility, we deployed the DMIC as an intelligent diagnostic device for multi-cancer classification (Fig. 5e). The system integrates a two-step decision architecture: a support vector machine (SVM) for initial cancer detection and a fully connected neural network (FCNN) for subsequent cancer-type discrimination. These models were programmed onto the scalable DMIC hardware, where clinical miRNA inputs were processed via diagnostic molecular chiplets navigating a unified fluidic

von Neumann architecture. (Figs. S44-50). For clinical validation, miRNA extracted from 1 mL of clinical serum was off-chip amplified and transformed into long single-stranded DNA (ssDNA) to preserve the original quantity ratios for on-chip computation (Figs. S51 and S52). The on-chip molecular classifier was validated first using 20 synthetic samples, achieving 100% classification accuracy (Fig. S53), and then tested on clinical serum samples from 15 cancer patients with lung (LC), liver (HCC), and ovarian (OC) cancers and 5 healthy controls, showing 90.0% sensitivity, 96.7% specificity, and 90.0% correct cancer-type prediction in positive cases (Figs. 5f,g and Fig. S54). These results highlight the platform's capability for automated sample-to-answer analysis, positioning the DMIC as a powerful and programmable tool for next-generation point-of-care diagnostics and integrated molecular intelligence.

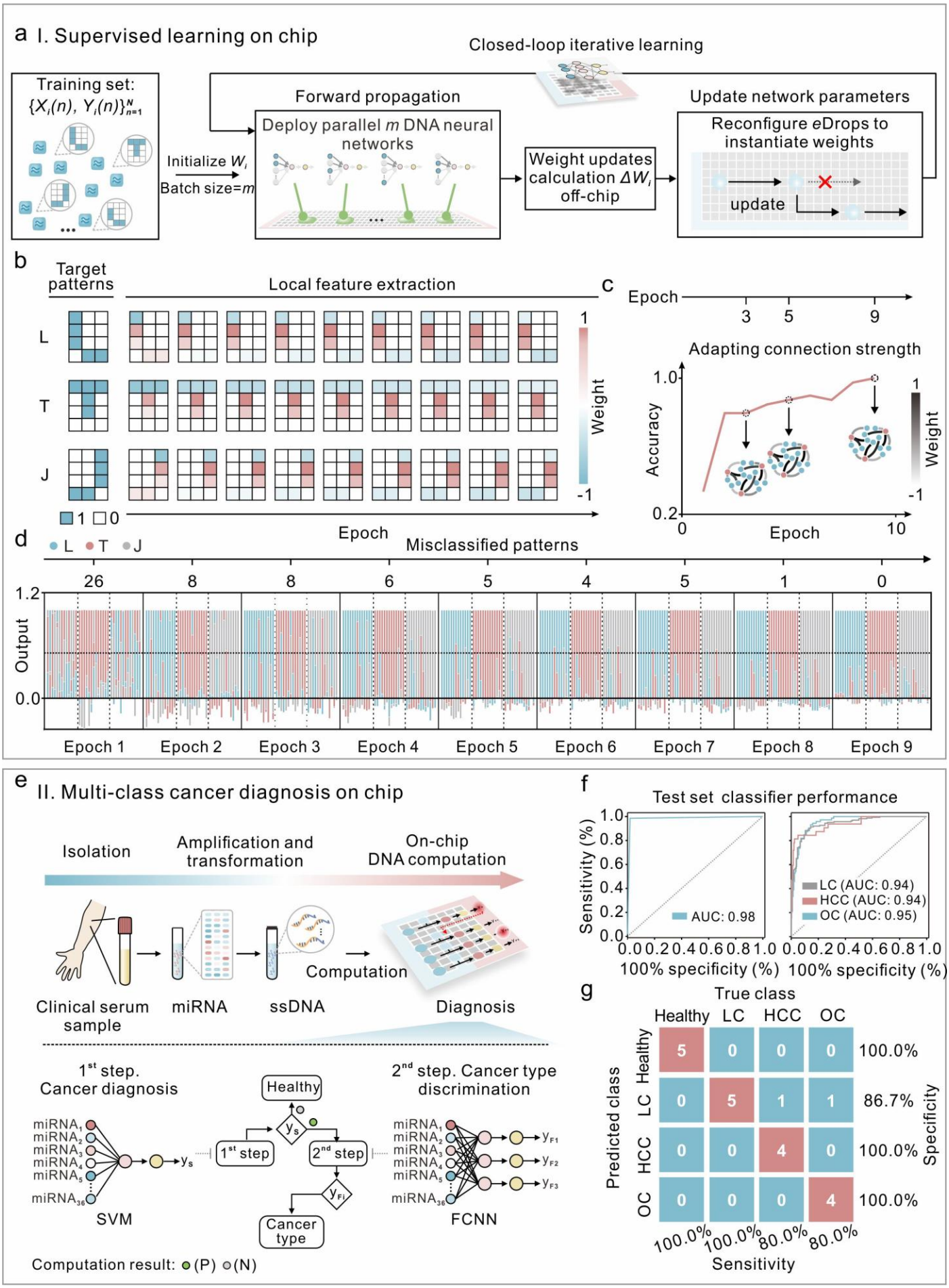


Fig. 5| Adaptive molecular intelligence through on-chip supervised learning and clinical multi-cancer diagnosis. a-d,

On-chip supervised learning via physical weight updates. **a,** Schematic of the closed-loop iterative learning process. Parallel DNA neural networks execute forward propagation on the integrated fluidic circuit, while weight updates (ΔW_j) calculated off-chip are used to dynamically and programmatically reconfigure the molecular chiplets for subsequent epochs. **b,** Progressive emergence of discriminative local features for patterns 'L', 'T', and 'J' as kernel weights physically evolve across training cycles. **c, d,** Accuracy curve (**c**) and the detailed evolution of network outputs (**d**) showing the adaptation of connection strengths during training. The classification task was considered completed when each output signal y_i corresponded exclusively to the correct class of the applied pattern. **e-g,** Intelligent multi-cancer diagnosis using clinical samples. **e,** Schematic showing the end-to-end diagnostic workflow from clinical serum extraction to on-chip DNA computation. The system integrates a two-step decision process (an SVM for cancer diagnosis and an FCNN for cancer-type discrimination) mapped onto a unified fluidic von Neumann architecture. In each step, each selected miRNA is assigned an integer-valued weight reflecting its significance. A negative output y_s indicates a healthy sample; a positive output y_s along with a positive output y_{Fi} leads to the identification of the cancer type (LC, HCC, or OC). **f,** Receiver operating characteristic (ROC) curves illustrating diagnostic performance on test sets. Classifier models were trained using publicly available miRNA profiles from the NCBI database. The area under the ROC curve (AUC) shown in each graph, is simply defined as the area under the ROC curve. **g,** Confusion matrix analysis of 20 clinical serum samples.

Discussion

The DMIC presented here overcomes the physical bottlenecks intrinsic to bulk-solution molecular computing systems by engineering a structured, programmable integrated fluidic circuit. We establish a fluidic von Neumann architecture that decouples the molecular computing substrate (molecular chiplets) from its control logic (electronic programming). Encapsulating DNA computational elements into discrete droplet-based molecular chiplets eliminates off-target molecular crosstalk. Addressable manipulation of these chiplets via an electrode array allows us to physically wire computing networks on-demand, transforming molecular computation from a diffusion-limited, equilibrium process into a programmable non-equilibrium system.

This architecture introduces crucial modularity and programmability. A library of standardized molecular chiplets functions as the system's instruction set, while the electrode grid corresponds to a reconfigurable global interconnect. In modern electronics, the chiplet architecture has revolutionized the semiconductor industry by decoupling system complexity from monolithic manufacturing limits, utilizing advanced interconnects to integrate modular components into highly scalable systems. Analogously, our DMIC platform decouples molecular computation from the physical constraints of homogeneous mixtures. By utilizing the electrode array to dynamically 'wire' standardized molecular chiplets, simply reprogramming droplet

pathways allows diverse network topologies to be instantiated on the same hardware. This integrated fluidic circuit approach dramatically simplifies design, enables resource reuse, and facilitates massively parallel computation, ultimately endowing molecular networks with unprecedented scalability. As a hybrid bioelectronic platform, DMIC uniquely suits intelligent, autonomous biosensing for integrated sample-to-answer analysis, opening avenues for next-generation point-of-care diagnostics, environmental monitors, and smart interfaces for synthetic biology.

Compared to conventional liquid-handling robots, the DMIC introduces a fundamental architectural leap in programmability and scalability. While commercial robots are restricted to bulk reagent distribution, they lack the capacity for the high-density, individualized orchestration of molecular chiplets required for complex network operations[44,45]. In contrast, our platform acts as a physical 'molecular compiler'—utilizing per-droplet addressing and dynamic rerouting to physically instantiate sophisticated, non-fully-connected neural topologies on demand. Looking forward, the computational capacity of the DMIC will follow a hardware-driven scaling law. Because physical scalability is now primarily constrained by active-matrix electrode density rather than chemical crosstalk, advanced microfabrication provides a direct pathway to massive network expansion. Furthermore, while the current architecture masterfully achieves topological reconfiguration via droplet routing, the in situ modification of internal molecular contents after droplet fusion remains a challenge. Future integration of auxiliary converter droplets for continuous on-chip parameter updating could further unlock fully autonomous network evolution. Ultimately, we envision that coupling this integrated fluidic circuit with automated nucleic acid extraction and amplification modules[46,47] will establish closed-loop, sample-to-answer systems, paving the way for scalable, embedded molecular intelligence.

Data and materials availability: The associated data and experimental protocols for this work are available within Supplementary Information, and are also available from the corresponding author on request. Supplementary Information is available in the online version of the paper. The source code for generating the droplet routing paths used in this work is available on GitHub at https://github.com/vivipurplr/DMIC-Droplet-Routing.

Acknowledgments: This work was supported by the National Key R&D Program of China (2021YFF1200300), the National Science Foundation of China (Grant No. T2188102, 21991134), the Shanghai Science and Technology Committee (STCSM) (23ZR1479600), Shanghai Pilot Program for Basic Research, Shanghai Municipal Science and Technology Major Project, and the New Cornerstone Science Foundation.

Author contributions: H.P. and C.F. conceived and supervised the project. M.C. and X.X. designed and performed the experiments. H.P., C.F., Q.Y., M.C., and X.X. discussed the design. M.C., X.X., J.W., S.X., N.L., and Y.Z. carried out experiments and interpreted data. B.L., T.Z., C.C., and N.L. developed the model and performed the *in-silico* training. S.H., H.M., L.X., and J.J. designed and fabricated the chip. All authors contributed to data analysis. X.X., L.L., M.C., F.W., C.F., and H.P. wrote the manuscript.

Competing interests: All authors declare that they have no competing interests.

References

1. Lewis, J. E. & Kemp, M. L. Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. *Nat. Commun.* **12**, 2700 (2021).
2. Wang, H. *et al.* Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47-60 (2023).
3. Sammut, SJ. *et al.* Multi-omic machine learning predictor of breast cancer therapy response. *Nature* **601**, 623-629 (2022).
4. Osipov, A. *et al.* The Molecular Twin artificial-intelligence platform integrates multi-omic data to predict outcomes for pancreatic adenocarcinoma patients. *Nat. Cancer* **5**, 299-314 (2024).
5. Thomasian, N. M., Kamel, I. R. & Bai, H. X. Machine intelligence in non-invasive endocrine cancer diagnostics. *Nat. Rev. Endocrinol.* **18**, 81-95 (2022).
6. Yang, S. *et al.* DNA as a universal chemical substrate for computing and data storage. *Nat. Rev. Chem.* **8**, 179-194 (2024).
7. Kieffer, C., Genot, A. J., Rondelez, Y. & Gines, G. Molecular computation for molecular classification. *Adv. Biology* **7**, e2200203 (2023).
8. Nagipogu, R. T., Fu, D. & Reif, J. H. A survey on molecular-scale learning systems with relevance to DNA computing. *Nanoscale* **15**, 7676-7694 (2023).
9. Hjelmfeltt, A., Weinbergert, E. D. & Ross, J. Chemical implementation of neural networks and turing machines. *Proc. Natl. Acad. Sci.* **88**, 10983-10987 (1991).
10. Li, J., Green, A. A., Yan, H. & Fan, C. Engineering nucleic acid structures for programmable molecular circuitry and intracellular biocomputation. *Nat. Chem.* **9**, 1056-1067 (2017).
11. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585-1588 (2006).
12. Gines, G. *et al.* Isothermal digital detection of microRNAs using background-free molecular circuit. *Sci. Adv.* **6**, eaay5952 (2020).
13. Xiao, M. *et al.* Boosting selective fusion of protocells with DNA logic circuits for in situ detection of exosomal microRNA. *Chem* **10**, 3634-3643 (2024).
14. Sun, J. *et al.* Implementing complex nucleic acid circuits in living cells. *Sci. Adv.* **11**, eadv6512 (2025).
15. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368-372 (2011).
16. Cherry, K. M. & Qian, L. Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* **559**, 370-376 (2018).
17. Xiong, X. *et al.* Molecular convolutional neural networks with DNA regulatory circuits. *Nat. Mach. Intell.* **4**, 625-635 (2022).
18. Okumura, S. *et al.* Nonlinear decision-making with enzymatic neural networks. *Nature* **610**, 496-501 (2022).
19. Lobato-Dauzier, N. *et al.* Neural coding of temperature with a DNA-based spiking chemical neuron. *Nat. Chem. Eng.* **1**, 510-521 (2024).
20. Song, T. & Qian, L. Heat-rechargeable computation in DNA logic circuits and neural networks. *Nature* **646**, 315–322 (2025).

21. Cherry, K. M. & Qian, L. Supervised learning in DNA neural networks. *Nature* **645**, 639–647 (2025).
22. Lopez, R., Wang, R. f. & Seelig, G. A molecular multi-gene classifier for disease diagnostics. *Nat. Chem.* **10**, 746-754 (2018).
23. Zhang, C. *et al.* Cancer diagnosis with DNA molecular computation. *Nat. Nanotechnol.* **15**, 709-715 (2020).
24. Ma, Q. *et al.* An automated DNA computing platform for rapid etiological diagnostics. *Sci. Adv.* **8**, eade0453 (2022).
25. Yin, F. *et al.* DNA-framework-based multidimensional molecular classifiers for cancer diagnosis. *Nat. Nanotechnol.* **18**, 677-686 (2023).
26. Yang, L. *et al.* A spatially localized DNA linear classifier for cancer diagnosis. *Nat. Commun.* **15**, 4583 (2024).
27. Chen, YJ., Groves, B., Muscat, R. A. & Seelig, G. DNA nanotechnology from the test tube to the cell. *Nat. Nanotechnol.* **10**, 748-760 (2015).
28. Qian, L. & Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**, 1196-1201 (2011).
29. Wang, F. *et al.* Implementing digital computing with DNA-based switching circuits. *Nat. Commun.* **11**, 121 (2020).
30. Lv, H. *et al.* DNA-based programmable gate arrays for general-purpose DNA computing. *Nature* **622**, 292-300 (2023).
31. Xiong, X. *et al.* Optochemical control of DNA-switching circuits for logic and probabilistic computation. *Angew. Chem. Int. Ed.* **60**, 3397-3401 (2021).
32. Tang, Q. *et al.* Multi-mode reconfigurable DNA-based chemical reaction circuits for soft matter computing and control. *Angew. Chem. Int. Ed.* **60**, 15013-15019 (2021).
33. Zhu, Y. *et al.* Accelerating DNA computing via freeze-thaw cycling. *Sci. Adv.* **9**, eaax7983 (2023).
34. Benenson, Y. Biomolecular computing systems: principles, progress and potential. *Nat. Rev. Genet.* **13**, 455-468 (2012).
35. Zhang, D. Y. & Winfree, E. Control of DNA strand displacement kinetics using toehold exchange. *J. Am. Chem. Soc.* **131**, 17303–17314 (2009).
36. Reif, J. H. Scaling up DNA computation. *Science* **332** (2011).
37. Sawai, S., Thomason, P. A. & Cox, E. C. An autoregulatory circuit for long-range self-organization in Dictyostelium cell populations. *Nature* **433**, 323–326 (2005).
38. Karsenti, E. Self-organization in cell biology: a brief history. *Nat. Rev. Mol. Cell Biol.* **9**, 255–262 (2008).
39. Dudukovic, N. A. *et al.* Cellular fluidics. *Nature* **595**, 58-65 (2021).
40. Finc, K. *et al.* Dynamic reconfiguration of functional brain networks during working memory training. *Nat. Commun.* **11**, 2435 (2020).
41. Wang, D. *et al.* Active-matrix digital microfluidics for high-throughput, precise droplet handling. *Nat. Rev. Electr. Eng.* 1-15 (2025).
42. Jia, Z. et al. Intelligent single-cell manipulation: LLMs- and object detection-enhanced active-matrix digital microfluidics. *Microsyst. Nanoeng.* **11**, 133 (2025).
43. Zhou, X. Understanding the convolutional neural networks with gradient descent and backpropagation. *J. Phys.: Conf. Ser.* **1004**, 012028 (2018).
44. Holland, I. & Davies, J. A. Automation in the Life Science Research Laboratory. Front. Bioeng. Biotechnol. 8, 1326 (2020).
45. Kong, F., Yuan, L., Zheng, Y. F. & Chen, W. Automatic liquid handling for life science: a critical review of the current state of the art. J. Lab. Autom. 17, 169–185 (2012).
46. Lin, H. *et al.* Ferrobotic swarms enable accessible and adaptable automated viral testing. *Nature* **611**, 570-577 (2022).
47. Ruan, Q. *et al.* Digital-WGS: Automated, highly efficient whole-genome sequencing of single cells by digital microfluidics. *Sci. Adv.* **6**, eabd6454 (2020).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.pdf