applied system innovation

MDPI

Communication

An Electro-Optic Lithium Niobate on Insulator (LNOI) Multimode Interference Accelerator for Reconfigurable Optical Neural Networks

Rong-Rong Nien * and Han-Chao Chang

National Center for Instrumentation Research, National Institutes of Applied Research, Hsinchu 300092, Taiwan; roman@niar.org.tw
* Correspondence: 2309248@niar.org.tw

Abstract

To overcome the thermal loss and interconnect delays of conventional electronic processors under escalating AI computing workloads, analog optical processors have emerged as a compelling low-latency alternative. This work introduces an integrated optical neural accelerator leveraging an electro-optically reconfigurable multimode interference (MMI) framework on a lithium-niobate-on-insulator (LNOI) platform. By harnessing localized spatial mode interference through micro-electrodes strategically placed within the MMI region, the device performs trainable matrix operations with a minimal on-chip insertion loss of 1.42 dB. A hardware-in-the-loop in situ (HIL in situ) calibration framework employing the Nelder–Mead optimization algorithm is developed to adaptively adjust local refractive indices and compensate for physical hardware imperfections. The system's performance is validated through numerical electromagnetic simulations applied to the famous Iris flower classification benchmark, demonstrating the self-learning capabilities of the integrated optical neural network with an accuracy of 72%. These findings present a compact, low-loss building block for next-generation photonic tensor processors.

Keywords: optical neural network; multimode interference; lithium niobate on insulator; hardware-in-the-loop in situ; photonic tensor processor

Check for updates

Academic Editor: Abdelkader Sbihi

Received: 28 August 2026
Revised: 15 October 2026
Accepted: 23 October 2026
Published: 27 October 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the International Institute of Knowledge Innovation and Invention. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Artificial intelligence (AI) is increasingly integrated into everyday life, necessitating substantial computational resources to manage the growing volume of data. In response, non-von Neumann computing architectures have been developed. For example, task-specific hardware accelerators such as graphics processing units (GPUs) [1,2] and tensor processing units (TPUs) [3] represent a novel computational paradigm compared to conventional computers. Nevertheless, during the transport of charged particles, electrical energy is dissipated as heat due to the electronic transmission mechanism. Electronic interconnects are subject to RC parasitic effects and electromigration, which contribute to power-intensive device operation [4–6]. In contrast, optical computation, wherein light propagates as an instantaneous resonance of wave attributes rather than through particle collisions, can mitigate heat loss and address challenges related to thermal dissipation.

Furthermore, wavelength-division multiplexing (WDM) technology enables the simultaneous transmission of multiple independent signals without crosstalk, analogous to parallel tensor operations on objects with distinct characteristics [7,8]. Beyond wavelength, optical properties such as polarization state, incidence angle, intensity, phase, and modal

propagation can also encode distinct characteristics. These properties impart unique values to each beam, facilitating phenomena such as beam interference, holographic recording, and electro-optic or thermo-optic modulation, which are instrumental in training optical neural networks. Photonic tensor processors commonly utilize phase-change materials (PCMs) in conjunction with Mach–Zehnder interferometer (MZI) meshes [4,9–11]. However, PCMs are constrained by binary absorption contrast and limited endurance, while MZI meshes require substantial physical footprints (exceeding 1 mm^2) and experience cumulative insertion losses in multi-port configurations [4]. Beyond PCM-based architectures, advanced material platforms exploit intrinsic defect engineering and optimization of physical parameters to realize dynamic optical modulation. For instance, in functional halide perovskites and perovskite manganites, manipulation of defect states, lattice phase optimization, and multi-physics modeling provide a robust framework for tuning local optoelectronic properties and addressing high-dimensional dynamic modeling and hardware optimization challenges [12–14]. A similar material engineering strategy applies to functional ferroelectric crystals such as lithium niobate ($LiNbO_3$) [15–19]. Within the lattice structure of pure $LiNbO_3$, niobium antisite defects (Nb_{Li}^{4+}) and lithium vacancies ($V_{Li}^{'}$) naturally form to compensate n-type carriers and maintain charge neutrality. Nevertheless, these deep trap sites often induce undesirable carrier accumulation and space-charge instabilities under high optical power. To mitigate these effects and enhance electro-optic responsiveness, magnesium oxide (MgO) doping is strategically introduced [20,21]. When the MgO concentration exceeds a critical threshold, Mg^{2+} ions substitute for niobium antisite vacancies, significantly reducing the density of deep traps and eliminating fluctuations in the built-in electric field. This defect-suppressed lattice markedly improves the material's resistance to photodamage, ensuring stable light propagation. More importantly, it enables a pure, high-speed linear electro-optic effect, whereby an externally applied electric field directly modulates the relative permittivity and extraordinary refractive index at sub-picosecond timescales without the limitations imposed by carrier transport.

Multimode interference (MMI) couplers facilitate direct multi-channel spatial mixing within a single compact cavity [22,23]. In the proposed device architecture, the MgO-doped LNOI layer functions as a highly stable, low-loss MMI waveguide, wherein integrated microelectrodes dynamically control spatial mode interference patterns to realize reconfigurable optical tensor computing. The resulting outputs can be fed back to adjust interference conditions or modulation parameters, culminating in the final optical neural network model. This methodology represents a significant advancement in optical neural network development [24]. Analogous to the backpropagation algorithm, optical modulation and storage directly influence the transmission parameters of subsequent input beams. Distinct from conventional platforms such as silicon-on-insulator (SOI) [25], silicon nitride [26], and indium phosphide on silicon [27], the present study employs a thin-film LNOI substrate. LNOI offers a high linear electro-optic coefficient, a broad transparency window, and strong dielectric optical confinement. Exploiting these unique physical properties, we propose a compact 4 × 3 reconfigurable electro-optic LNOI MMI tensor processor that achieves a low-total on-chip insertion loss of 1.42 dB. A detailed loss budget analysis confirms that this total loss is effectively managed across all contributing physical mechanisms, including intrinsic MMI self-imaging and mode-transition loss (0.85 dB) for millimeter scale [22,27,28], waveguide absorption loss (0.35 dB) [29], and taper/junction loss (0.22 dB) [30]. By consolidating spatial tensor operations within a single defect-suppressed LNOI cavity, this approach obviates the cascading insertion loss penalties characteristic of traditional MZI meshes, which multiply losses of 0.5–1.2 dB by N in N-cascaded topologies, then potentially totaling up to 3 dB [11]. Therefore, MMI-based architecture establishes a robust, low-loss architecture for dynamic optical computing.

In this work, we have designed the MMI device with multiple electrodes strategically placed within the interference region to leverage the electro-optic effect, thereby imparting self-learning capabilities to the device. The methodologies applied are comprehensively described in Section 2, while the results of numerical simulations are presented in Section 3. We have demonstrated a pre-trained inference task focused on the classification of Iris flower species. Subsequently, an analysis of the optical neural network inference is provided, followed by a discussion in the concluding section. For the pretrained model employed in this study, the initial refractive index was determined prior to training to facilitate preliminary validation. To contextualize the performance and architectural innovations of the proposed reconfigurable neural unit, a detailed comparison with recent on-chip optical neural networks (ONNs) utilizing multimode interference (MMI)-like structures is provided in Table 1. This comparison includes key parameters such as material composition, operational characteristics, and fundamental mechanisms.

Table 1. Comparison of the on-chip ONN using MMI-like structures.

Design	Substrate	Input/ Outputs	Trainable Parameters	Input Attribute	Modulated Mechanism	Training Scheme	Optical Loss [dB]	Iris Accuracy
Bai et al. (2023) [31]	SOI	10 × 10	20	wavelength	MRR	In situ GDC	4	97.0% (MNIST)
Fu et al. (2023) [25]	SOI	4 × 3	210	phase	SWU	External	6.5	90%
Meng et al. (2023) [26]	SiN	10 × 10	9	wavelength	phase shifter	External	1.5	92.17% (MNIST)
Lin et al. (2024) [32]	TFLN	10 × 10	300	wavelength	MZM	In situ	0.02	92.0% (MNIST)
Ghosh et al. (2025) [27]	InP on Si	4 × 3	3	wavelength	liquid crystal	External	<0.5	86.67%
This work	LNOI	4 × 3	3	amplitude	MMI	In situ (HIL)	1.42	72.0%

2. Materials and Methods

2.1. Device Design

Multimode interference finds extensive applications within the field of photonics [33–35], frequently employing materials such as SOI, silica, or indium phosphide (InP) [27,36] as the core medium for the multimode interference region. To realize a reconfigurable optical neural unit, an integrated multimode interference coupler embedded with a micro-electrode array was designed on a LNOI substrate. Recently, LNOI has also been utilized as the MMI core material [37,38], leveraging its electro-optic properties for modulation. The lithium niobate layer is etched to form a 4 × 3 MMI structure. The conceptual framework of the ONN device is depicted in Figure 1a. Corresponding to the deep learning backpropagation process illustrated in Figure 1b, the primary function of this device is the reverse allocation of errors and precise attribution [39]. Fundamentally, this mechanism quantifies the contribution of each component to the final error, thereby enabling the dynamic replacement of ineffective features and the refinement of convolutional layers' capacity to extract critical edges, textures, and higher-order semantic features. This process constitutes the core operational engine that endows the neural network with self-correction and adaptive learning capabilities.

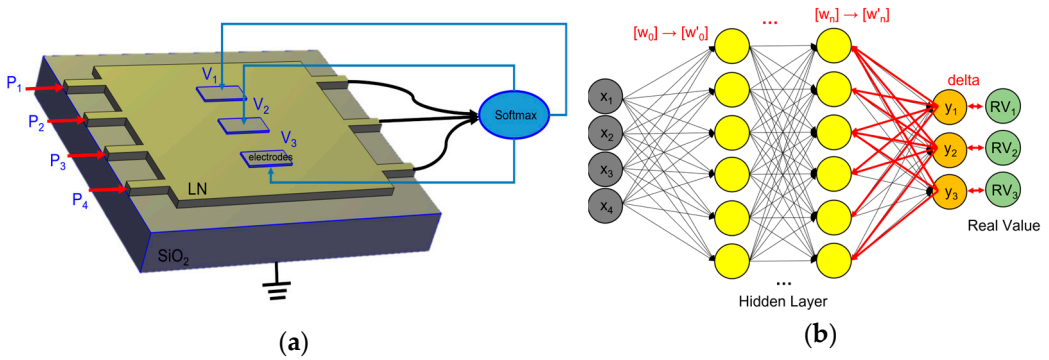


Figure 1. Diagrammatic representation of the reconfigurable optical neural unit: (**a**) cross-sectional arrangement of the thin-film lithium-niobate-on-insulator (LNOI) rib waveguide integrated with embedded micro-electrodes, and (**b**) the self-learning optimization framework emulating the information flow characteristic of neural network backpropagation.

Lithium niobate is an electro-optic crystal. When subjected to an applied electric field E, the change in its refractive index follows the relationship described by Equation (1):

$$\Delta n = -\frac{1}{2} r n^3 E. \tag{1}$$

In this equation, r is referred to as the electro-optic coefficient. Following the inference process of the deep learning model, the optical power detected at the three output ports can be amplified through an external circuit and subsequently converted into an output voltage. The fundamental architectural design comprises a standardized multimode interference mixing region, four input single-mode waveguides utilized for feature phase encoding, and three dedicated output ports that function as structural neurons during inference (see Figure 1a). The schematic configuration and the multi-physics coupling mechanism of the proposed device are presented herein.

The beat length for the definition of the MMI is defined as:

$$L_\pi \approx \frac{4 n_{eff} W_{MMI}^2}{3\lambda_0}, \tag{2}$$

where n_{eff} denotes the effective refractive index, and W_{MMI} represents the width of the MMI region. The input mode pattern replicates itself over a distance corresponding to three times the beat length [35,36]. The finite element method (FEM) was employed to compute the output mode distribution. Due to computational limitations associated with scaling large multi-physics meshes, a compact MMI structure with an enhanced refractive index tuning range was utilized in this study to assess the topological self-learning capability. In practical thin-film lithium niobate (TFLN) applications, this equivalent phase modulation can be effectively implemented in a physically realistic setting by proportionally scaling the length of the electro-optic mixing region to the millimeter scale. This approach successfully preserves the wave-front steering functionality while avoiding dielectric breakdown. For further details, refer to Figure 2, which illustrates that the distance between the input and output ports is increased tenfold, while maintaining the same phase condition when the refractive index is decreased by an order of magnitude. To minimize the computational overhead associated with multi-physics meshing during iterative numerical solver cycles, a scaled numerical model was developed. The validity of spatial scaling is determined by phase conservation principles. The resulting accumulated modal phase shift is described by the following relation:

$$\phi = \frac{2\pi}{\lambda_0} \Delta n \cdot L_{MMI}. \tag{3}$$

By maintaining a constant product of the refractive index and device length ($\Delta n_1 \cdot L_1 = \Delta n_2 \cdot L_2$), scaling the physical length of the device while proportionally reducing the required index modulation preserves the optical phase accumulation precisely. Moreover, since $\Delta n \propto E_z$, the voltage-length product remains unchanged, thereby validating that compact numerical models accurately represent the behavior of physical millimeter-scale TFLN devices operating under CMOS-compatible driving voltages ($\pm 15V$). Calibrated wave propagation profiles mapped onto the physical MMI coordinates are illustrated in Figure 2b.

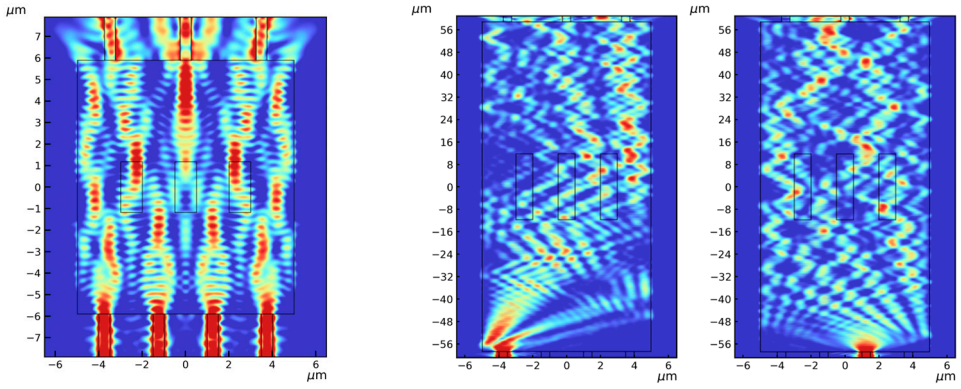


Figure 2. COMSOL multiphysics simulation illustrating the wave propagation characteristics and light-splitting profiles within the multimode interference (MMI) mixing region subjected to electro-optically modulated refractive index perturbations (Δn). (**a**) The propagation pattern within the MMI region, and (**b**) the output patterns observed at various ports over a tenfold propagation distance.

The optical accelerator transforms input optical power vectors $\mathbf{P}_{\text{in}} \in \mathbb{R}^4$ into output optical power vectors $\mathbf{P}_{\text{out}} \in \mathbb{R}^3$. To explicitly define the relationship between the electrical control parameters and the neural network weight matrix $\mathbf{W}$, the complex output electric field vector $\boldsymbol{E}_{out}$ is related to the input electric field vector $\boldsymbol{E}_{in}$ via a voltage-dependent complex optical transfer matrix $\mathbf{T}(\mathbf{V})$:

$$\boldsymbol{E}_{out} = \mathbf{T}(\mathbf{V})\boldsymbol{E}_{in}, \tag{4}$$

where $\mathbf{T}(\mathbf{V})$ is determined by the cascaded modal propagation matrices connecting the electrode regions and the localized electro-optic phase modulation profile. Specifically, each micro-electrode voltage V_j generates a vertical electric field $E_z^{(j)}$, which linearly modulates the local extraordinary refractive index according to Equation (1). This perturbation in the refractive index modifies the internal spatial phase distribution and mode coupling within the MMI cavity. Since the square-law photodetectors at the output ports measure the optical intensity ($P_{out,i} = |E_{out,i}|^2$), the corresponding neural weight transformation matrix $\mathbf{W}(\mathbf{V})$ is nonlinearly derived from the optical transfer matrix elements as $W_{ij} = |T_{ij}(\mathbf{V})|^2$. This establishes a precise mathematical relationship between the physical electro-optic modulation and the network weight configurations.

To quantify the independent control capability and spatial crosstalk among the three micro-electrode actuators, we introduce the sensitivity matrix $\mathbf{S} = \partial \mathbf{P}_{out}/\partial \mathbf{V}$ [40], which relates small voltage perturbations around the operating bias point to corresponding changes in output optical power:

$$\Delta \mathbf{P}_{out} = \mathbf{S}\Delta \mathbf{V}. \tag{5}$$

When subjected to a controlled voltage perturbation of ± 0.1 V/μm, the singular value decomposition of the matrix $\mathbf{S}$ reveals three non-zero singular values (σ_1,σ_2,σ_3), indicating the presence of three linearly independent degrees of parameter control. Moreover,

the sensitivity matrix demonstrates a notably low condition number $\kappa(\mathbf{S}) = \sigma_{\max}/\sigma_{\min} \approx 2.89$ and strong diagonal dominance. These characteristics ensure that each actuator independently controls distinct spatial output channels with minimal cross-coupling, thereby facilitating robust and non-singular gradient convergence throughout the HIL in situ optimization process.

The optical power values from the four input ports are utilized as variables for the input layer of the neural network. The operating wavelength employed is 1550 nm. Due to the significant variation among the port values, a linear correction is applied to the input optical power measurements. By testing various input values from the dataset, the simulated intensities at the three output ports are obtained for this Iris flower classification example. Subsequently, the Softmax function is applied to these intensities to derive probability outcomes, serving as the inference for each training epoch. The photodetector material may be either InGaAs or Ge. Upon detecting the light intensity, the photodetector converts this measurement into voltage parameters, which are then transmitted to a computer interfaced with the photodetector. These voltage values are subsequently relayed back to the micro-electrodes positioned on the MMI element. All geometric, material, and operational parameters are comprehensively detailed in Table 2. Our analysis indicates that the placement of the electrodes along the beat length direction exerts minimal influence on the distribution of the output patterns.

Table 2. Optimized physical and geometrical parameters of the 4 × 3 electro-optic LNOI MMI processor.

Category	Parameter	Symbol	Quantitative Value
Material Property	Substrate Orientation	—	Z-cut
	Electro-optic Coefficient	r	30.8 pm/V
	Effective Index	n_{eff}	2.19
	Extraordinary Index	n_e	2.13
Waveguide Geometry	MMI Region Width	W_{MMI}	10 μm
	Port Waveguide Width	W_{port}	500 nm
	MMI Region Length	L_{MMI}	566.4 μm ($3{\cdot}L_\pi$)
	Beat Length	L_π	188.8 μm
Simulation Setting	MMI Scaling Length	L_{MMI}	11.8 μm (Scaled down by 48)
Operating Condition	Wavelength	λ_0	1550 nm
	Applied Electric Field	E_z	3.36 V/μm
	Refractive Index Modulation	Δn	~5×10^{-4}

2.2. Data Processing

We develop a simulator using COMSOL Multiphysics 6.4 and Python 3.10. The solver for the reconfigurable optical neural unit is implemented using a multi-threaded Python framework to facilitate synchronous hardware-in-the-loop (HIL) in situ optimization [41], enabling both initialization and real-time performance visualization. The solution workflow is depicted in Figure 3. The primary thread is responsible for system initialization and data loading, employing a non-blocking loop to continuously monitor the shared is_updated() flag for incoming simulation results. The Iris flower dataset comprises N feature vectors. Simultaneously, a dedicated optimization thread establishes a live connection to the COMSOL multi-physics server, loads the specified LNOI model, and initializes the SciPy optimization routine. This optimization loop employs the Nelder–Mead simplex algorithm to identify the

optimal electro-optic tuning parameters across the three integrated electrodes. During each iteration, the get_mmi_loss() function is called, triggering a nested process that iterates over a batch of training data. Within this data processing workflow, each element of an Iris flower feature vector is encoded to correspond to optical power, which is sequentially applied to the COMSOL model along with the refractive index weights. The model is then solved, the optical power at the output ports is evaluated, and the cross-entropy loss is computed. The function returns the loss value to the optimizer while asynchronously updating the confusion matrix in shared memory. This process encompasses three primary components of data processing: confusion matrix visualization, the SciPy optimization routine, and the COMSOL data iteration thread.

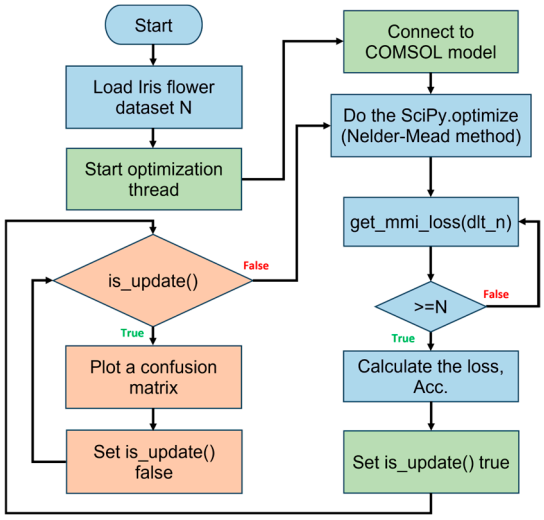


Figure 3. Operational flowchart illustrating the data preprocessing framework, specifically outlining the normalization and mapping processes that transform the input geometric features of the Iris dataset into dynamic phase-encoding profiles.

To evaluate the inherent generalization capability of the hardware with minimal computational expense, the Nelder–Mead optimization algorithm was applied using a training subset comprising only 3 representative samples, one from each class (Setosa, Versicolor, and Virginica). The voltages of the three micro-electrodes were directly optimized within the HIL setup to define the primary decision boundary separating the three categories in the optical field domain. Upon determining the three electrode control parameters, the fixed electro-optic MMI accelerator was tested on the complete Fisher's Iris flower dataset of 150 samples, thereby subjecting the system to previously unseen testing conditions for 147 samples. The device attained a testing accuracy of 72.0%. Given that only three control voltages were tuned on three training samples, the system is inherently precluded from parameter memorization or overfitting. Consequently, the observed 72.0% testing accuracy substantiates that the physical modal transformation within the electro-optic MMI intrinsically captures the generalized spatial feature distribution of the Iris dataset. Analysis of the confusion matrix reveals that Class 0 (Setosa) achieves 90% precision, attributable to pronounced optical field separation. Misclassifications predominantly occur between Class 1 (Versicolor) and Class 2 (Virginica), due to overlapping spatial features within the constrained three-parameter optical transformation space. Since the number of trainable parameters equals the number of training exemplars, overfitting is mathematically excluded, thereby confirming that the 72.0% accuracy genuinely reflects the physical feature mapping capability of the three-electrode MMI architecture.

3. Results

The refractive indices of lithium niobate at a wavelength of 1550 nm are reported as $n_o = 2.21$ for ordinary rays and $n_e = 2.13$ for extraordinary rays. Additionally, the real part of the effective refractive index, $n_{eff} = 2.19$, for LNOI was determined through the effective index method (EIM) implemented in COMSOL Multiphysics. Utilizing Equation (1) based on the electro-optic effect, the change in refractive index induced by the micro-electrodes was estimated. Feature vector values were encoded into exponential terms within an MMI framework and iteratively incorporated into COMSOL to generate confusion matrices and evaluate classification accuracies. The refractive index variation, Δn, was computed and optimized by employing the Nelder–Mead optimization method. For the Iris dataset comprising $N = 150$ samples, the classes setosa, versicolor, and virginica were evenly represented. The highest classification accuracy achieved was 69% using three micro-electrodes, corresponding to refractive index values ranging from 2.19 to 2.23 under an applied voltage range of 8.40 to 14.02 V, as illustrated in Figure 4a. Furthermore, an accuracy of 72% was attained with refractive index values between 2.17 and 2.20 under a voltage range of 5.60 to 9.81 V, as illustrated in Figure 4b. The overall optical loss budget and physical performance parameters were evaluated through simulations and theoretical LNOI waveguide models, as summarized in Table 3.

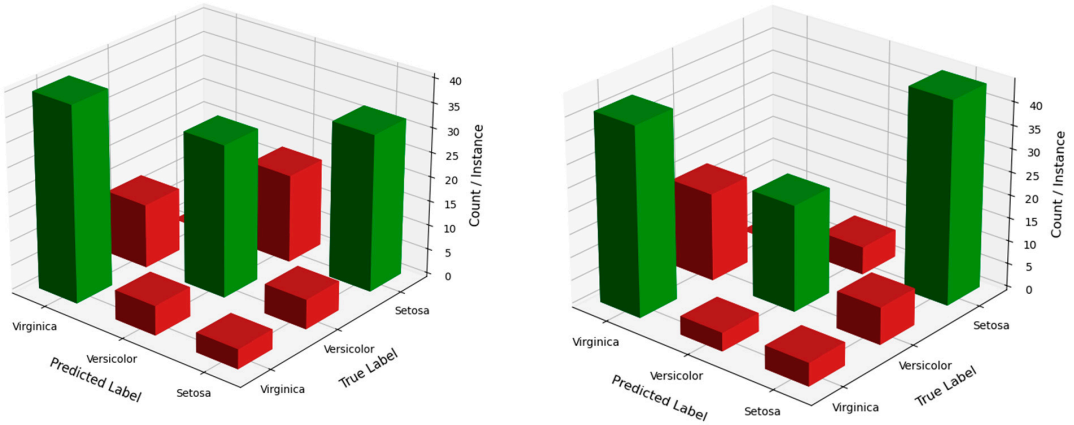


Figure 4. Displays the results of the three-dimensional confusion matrix optimization performed during the iterative optimization procedure for Iris flower classification. Panel (**a**) depicts the outcomes corresponding to a configuration with refractive index values between 2.19 and 2.23, whereas panel (**b**) illustrates the results for a configuration with refractive index values ranging from 2.17 to 2.20.

Table 3. Quantitative assessment of optical loss budget and operational performance parameters for the proposed LNOI MMI accelerator.

Parameter	Quantitative Value
Waveguide core pitch	370 nm
Waveguide rib height	260 nm
MMI insertion loss	0.85 dB
Waveguide absorption loss	0.35 dB
Taper/junction loss	0.22 dB
Total optical loss	1.42 dB
Driving voltage	3.36–14.02 V
Optical propagation latency	1.85 ps

4. Discussion

4.1. Simulation Framework Limitations

In the advancement of analog photonic computing and photonic tensor cores, the majority of state-of-the-art research predominantly depends on sophisticated phase encoding, multi-wavelength wavelength-division multiplexing (WDM) techniques, or dynamic liquid crystal modulation to attain exceptionally high classification accuracy on standard datasets [25–27,31,32]. In contrast, the platform presented in this study demonstrates a simulated classification accuracy of 72% on the Iris benchmark, representing an optimal architectural compromise achieved under minimal hardware constraints.

Our system employs straightforward intensity-based features for input encoding, thereby obviating the necessity for precise control over phase zeros and drift. This approach fundamentally overcomes the limitations associated with complex phase encoding and multi-wavelength encoding, substantially reducing both power consumption and the control complexity of the front-end optoelectronic interface. Nonetheless, current modeling and simulation frameworks exhibit several physical limitations: numerical simulations presume ideal sidewall smoothness and uniform material stoichiometry, whereas fabricated components frequently exhibit line-edge roughness, resulting in additional backscattering and spatial phase perturbations. Moreover, the framework models static electro-optic modulation but do not comprehensively capture transient thermal effects induced by dielectric loss. Additionally, discretization within the simulation mesh introduces minor phase dispersion errors over extended propagation distances, necessitating empirical calibration during experimental validation to precisely align the positions of self-imaging nodes.

4.2. Architectural Scalability

Accuracy can be further improved within a compact, single-stage footprint by increasing the spatial density of the micro-electrode array. This enhancement enriches the reconfigurable extraordinary refractive index profiles and matrix expressiveness without incurring cumulative insertion loss. Extending feature representations to complex phase–amplitude vectors and integrating optical nonlinear activation elements at the output ports will facilitate higher-dimensional nonlinear classification. In practical on-chip photonic fabrication, non-idealities such as line-edge roughness, material stoichiometry variations, backscattering, and dielectric-loss-induced thermal effects can cause significant spatial phase perturbations and wavelength drifts in high-precision phase-encoding systems. As a result, conventional architectures aiming for high classification accuracy (e.g., >90% accuracy) often require complex and costly real-time phase compensation and calibration mechanisms on physical hardware.

In contrast, the proposed architecture offers notable advantages in physical scalability and spatial compactness. Traditional optical tensor processors based on MZI mesh grids experience exponential growth in component count and dense waveguide routing as the channel number N increases, often resulting in footprints exceeding several square millimeters. Our single-order MMI design integrates optical interference and electro-optic modulation within a single compact cavity. With a waveguide width of approximately 10 μm and a total footprint below 0.1 mm^2, this structure allows for enhanced reconfigurable extraordinary refractive index distributions and matrix representation capacity by increasing the spatial density of the microelectrode array, thereby avoiding proportional physical expansion and cumulative insertion loss. Moreover, the platform exhibits exceptional energy efficiency. Due to the intrinsic capacitive characteristics of the TFLN electro-optic modulator, static leakage currents are negligible, resulting in near-zero static power consumption on the order of nanowatts. Dynamic energy consumption is determined by the charging and discharging behavior of the microelectrodes, quantified by the relation $E_{\text{dynamic}} = \frac{1}{2}CV^2$.

Given the ultra-low component capacitance in the femtofarad range and modest driving voltages below a few volts, the energy consumption per operation is rigorously maintained at an exceptionally low level of less than 1.5 pJ/OP.

The physical scalability of the proposed electro-optic MMI accelerator is determined by the invariance of the electro-optic figure of merit and the fidelity of the spatial phase field. As the MMI footprint expands to accommodate higher-port configurations, the required phase shift per unit length remains constant. To assess the mathematical stability during multi-channel voltage tuning, a condition number analysis was conducted on the theoretical sensitivity matrix as defined in Equation (5). The matrix demonstrates a notably low condition number, indicating that the mapping between micro-electrode voltage perturbations and spatial interference power variations is well-conditioned and far from singular. This low condition number ensures robust gradient convergence during HIL in situ optimization, avoiding instabilities associated with ill-posed matrix inversion.

Owing to the capacitive nature of LNOI electro-optic modulators, static power consumption is effectively negligible. Dynamic energy consumption is governed by electrode charging, enabling ultra-low power operation that is well-suited for energy-efficient edge AI processing. Moreover, long-term operational stability under high optical power densities is ensured through physical defect engineering of the LNOI platform. Specifically, the incorporation of magnesium oxide (MgO) doping above its critical threshold significantly reduces the density of deep-trap states and suppresses the accumulation of photorefractive space-charge fields, thereby preventing optical damage and drift.

4.3. Future Outlook

In summary, in contrast to the work of Meng et al. [26] and Ghosh et al. [27] and other researchers who have pursued highly complex on-chip optical systems to achieve superior performance, this study demonstrates a novel approach to "efficient analog photonic tensor computation with minimal intensity coding," validated by an Iris flower classification accuracy of 72%. The principal advantages of this approach include a substantial simplification of front-end optical hardware, exceptional architectural compactness and computational speed, as well as reduced manufacturing and calibration complexity with enhanced fault tolerance. This presents a highly valuable solution for the future large-scale integration of photonic computing processors into high-density optoelectronic hybrid chips.

Looking forward, the research focus will progressively shift from the current global optimization phase toward the development of a fully autonomous on-chip in situ training methodology. Within integrated photonics, "in-situ training" typically denotes fully autonomous on-chip backpropagation; however, in the present work, it is currently implemented as a minimal sample optimization technique. The system employs a physical LNOI MMI cavity for forward propagation to generate real-time optical intensity patterns, which are subsequently compared with target labels to update the microelectrode driving voltages. This closed-loop framework effectively compensates for unmodeled physical non-idealities—such as manufacturing tolerances, thermal drift, and waveguide sidewall scattering—thereby bridging the gap encountered when transferring offline digital training weights to actual hardware.

At present, the system depends on an external algorithmic solver to iteratively determine the optimal refractive index parameters. The long-term objective is to eliminate reliance on desktop computing by enabling error gradient backpropagation directly within the physical electro-optic hardware. By routing electrical error feedback to on-chip control circuitry, the LNOI MMI network would autonomously adjust microelectrode voltage settings in real time. This progression from assisted optimization to a fully independent, chip-level adaptive learning system is expected to significantly reduce control latency, improve ro-

bustness against local manufacturing defects, and facilitate the development of scalable, self-contained photonic deep learning processors suitable for edge computing applications.

5. Conclusions

This study presents a compact and energy-efficient analog photonic tensor processor built on a single-stage LNOI multimode interference cavity. Utilizing a simple intensity-based input encoding approach, the design eliminates the need for complex phase-tracking systems, significantly reducing front-end hardware complexity and power consumption. The device seamlessly integrates optical interference and electro-optic modulation within an ultra-compact footprint of less than 0.1 mm^2, achieving a simulated classification accuracy of 72% on the Iris dataset. Thanks to the inherent capacitive properties of thin-film lithium niobate, static leakage is minimal, keeping dynamic energy consumption below 1.5 pJ/OP. Additionally, condition number analysis verifies a well-conditioned sensitivity matrix, supporting reliable gradient convergence during hardware-in-the-loop in situ optimization. Key non-idealities, such as deep-trap state accumulation and photorefractive damage at high optical powers, are effectively addressed through optimal magnesium oxide doping.

While these results are promising, current numerical models have limitations in accurately capturing transient thermal effects and spatial phase disturbances caused by line-edge roughness and material non-uniformities. Future developments will focus on overcoming these challenges by increasing the spatial density of micro-electrodes, adopting more sophisticated phase–amplitude modeling, and incorporating optical nonlinear activation elements to support higher-dimensional operations. Importantly, upcoming research aims to shift from assisted optimization methods to fully autonomous on-chip backpropagation. By directly feeding electrical error signals into the control circuitry, the self-contained MMI network will dynamically compensate for manufacturing variations, thermal fluctuations, and scattering losses in real time. Overall, this architecture offers a highly scalable, fault-tolerant, and energy-efficient solution, paving the way for large-scale optoelectronic hybrid computing and edge AI processors.

Author Contributions: Software, R.-R.N.; validation, R.-R.N.; writing—original draft preparation, R.-R.N.; writing—review and editing, H.-C.C.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The structural geometries, simulation models, and optimization datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MMI	Multimode Interference
LNOI	Lithium Niobate on Insulator
HIL	Hardware-in-the-Loop
AI	Artificial Intelligence
GPU	Graphics Processing Units
TPU	Tensor Processing Units
WDM	Wavelength-Division Multiplexing

PCM	Phase-Change Materials
MZI	Mach–Zehnder Interferometers
SOI	Silicon on Insulator
ONN	Optical Neural Networks
MRR	Microring Resonator
SWU	Subwavelength Unit
MZM	Mach–Zehnder Modulator
MNIST	Modified National Institute of Standards and Technology
FEM	Finite Element Method
TFLN	Thin-Film Lithium-Niobate
CMOS	Complementary Metal–Oxide–Semiconductor
EIM	Effective Index Method

References

1. Chen, Y.-H.; Krishna, T.; Emer, J.S.; Sze, V. Eyeriss: An Energy-Efficient Reconfigurable Accelerator for Deep Convolutional Neural Networks. *IEEE J. Solid-State Circuits* **2017**, *52*, 127–138. [CrossRef]
2. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
3. Jouppi, N.P.; Young, C.; Patil, N.; Patterson, D.; Agrawal, G.; Bajwa, R.; Bates, S.; Bhatia, S.; Boden, N.; Borchers, A.; et al. In-Datacenter Performance Analysis of a Tensor Processing Unit. In *Proceedings of the 44th Annual International Symposium on Computer Architecture, Toronto, ON, Canada, 24–28 June 2017*; ACM: New York, NY, USA, 2017; pp. 1–12.
4. Shen, Y.; Harris, N.C.; Skirlo, S.; Prabhu, M.; Baehr-Jones, T.; Hochberg, M.; Sun, X.; Zhao, S.; Larochelle, H.; Englund, D.; et al. Deep Learning with Coherent Nanophotonic Circuits. *Nat. Photonics* **2017**, *11*, 441–446. [CrossRef]
5. Li, J.; Meng, X.; Zhang, J.; Li, W.; Shi, N.; Li, M. High-Efficiency Reconfigurable Optical Convolutional Neural Network. *J. Light. Technol.* **2025**, *43*, 4076–4085. [CrossRef]
6. Shi, B.; Calabretta, N.; Stabile, R. InP Photonic Integrated Multi-Layer Neural Networks: Architecture and Performance Analysis. *APL Photonics* **2022**, *7*, 010801. [CrossRef]
7. Shi, B.; Calabretta, N.; Stabile, R. Deep Neural Network Through an InP SOA-Based Photonic Integrated Cross-Connect. *IEEE J. Sel. Top. Quantum Electron.* **2020**, *26*, 7701111. [CrossRef]
8. Fu, T.; Zhang, J.; Sun, R.; Huang, Y.; Xu, W.; Yang, S.; Zhu, Z.; Chen, H. Optical Neural Networks: Progress and Challenges. *Light. Sci. Appl.* **2024**, *13*, 263. [CrossRef] [PubMed]
9. Xu, X.; Tan, M.; Corcoran, B.; Wu, J.; Boes, A.; Nguyen, T.G.; Chu, S.T.; Little, B.E.; Hicks, D.G.; Morandotti, R.; et al. 11 TOPS Photonic Convolutional Accelerator for Optical Neural Networks. *Nature* **2021**, *589*, 44–51. [CrossRef] [PubMed]
10. Feldmann, J.; Youngblood, N.; Wright, C.D.; Bhaskaran, H.; Pernice, W.H.P. All-Optical Spiking Neurosynaptic Networks with Self-Learning Capabilities. *Nature* **2019**, *569*, 208–214. [CrossRef] [PubMed]
11. Shokraneh, F.; Geoffroy-Gagnon, S.; Nezami, M.S.; Liboiron-Ladouceur, O. A Single Layer Neural Network Implemented by a 4 × 4 MZI-Based Optical Processor. *IEEE Photonics J.* **2019**, *11*, 4501612. [CrossRef]
12. Nurbekova, M.; Mataev, M.; Abdraimova, M.; Tursyn, Z.; Durmenbayeva, Z.; Sarsenbaeva, Z. Perovskite Manganites: An Overview of Synthesis, Classification, Characterization, and Applications. *IJMS* **2026**, *27*, 5709. [CrossRef] [PubMed]
13. Jabarov, S.H.; Hakhverdiyeva, Z.E.; Aliyev, Y.I.; Meshalkin, A.Y.; Trukhanov, S.V.; Trukhanov, A.V.; Huseynov, E.M. DSC Spectroscopic Investigation of Nd0.5Ca0.5MnO3, Pr0.5Ca0.5MnO3, and Nd0.5Sr0.5MnO3 Compounds. *Surf. Engin. Appl. Electrochem.* **2026**, *62*, 241–246. [CrossRef]
14. Jabarov, S.H.; Hakhverdiyeva, Z.E.; Samadov, S.F.; Sidorin, A.A.; Trung, N.V.M.; Abiyev, A.S.; Aliyev, Y.I.; Tiep, N.V.; Trukhanov, S.V.; Trukhanov, A.V. A-Site Cation Substitution Effects on Lattice Distortion and Open-Volume Defects in Nd–Ca/Sr and Pr–Ca Perovskite Manganites. *Micro Nanostructures* **2026**, *214*, 208659. [CrossRef]
15. Inbar, I.; Cohen, R.E. Origin of Ferroelectricity in $LiNbO_3$ and $LiTaO_3$. *Ferroelectrics* **1997**, *194*, 83–95. [CrossRef]
16. Hossain, M.M. First-Principles Study on the Structural, Elastic, Electronic and Optical Properties of $LiNbO_3$. *Heliyon* **2019**, *5*, e01436. [CrossRef] [PubMed]
17. Peng, L.; Hong, L.; Chen, B.; He, P.; Li, Z. Robust Modal Phase Matching in Subwavelength X-Cut and z-Cut Lithium Niobate Thin-Film Waveguides. *Chin. Opt. Lett.* **2021**, *19*, 071902. [CrossRef]
18. Zhu, D.; Shao, L.; Yu, M.; Cheng, R.; Desiatov, B.; Xin, C.J.; Hu, Y.; Holzgrafe, J.; Ghosh, S.; Shams-Ansari, A.; et al. Integrated Photonics on Thin-Film Lithium Niobate. *Adv. Opt. Photon.* **2021**, *13*, 242. [CrossRef]
19. Wang, L.; Du, H.; Zhang, X.; Chen, F. Optical Nonlinearity of Thin Film Lithium Niobate: Devices and Recent Progress. *J. Phys. D Appl. Phys.* **2025**, *58*, 023001. [CrossRef]

20. Schlarb, U.; Betzler, K. Influence of the Defect Structure on the Refractive Indices of Undoped and Mg-Doped Lithium Niobate. *Phys. Rev. B* **1994**, *50*, 751–757. [CrossRef] [PubMed]
21. Liu, B.-W.; Han, Z.-Q.-Z.; Li, J.-P.; Wang, X.-H.; Li, Y.-H.; Zhou, Z.-Y.; Shi, B.-S. Sellmeier Equation for the Refractive Index of Ordinary Light in 5 Mol.% MgO-Doped Congruent Lithium Niobate. *Opt. Express* **2025**, *33*, 1163. [CrossRef] [PubMed]
22. Le, T.-T.; Le, D.-T. Multimode Waveguides on an SOI Platform for Arbitrary Power Splitting Ratio Couplers. In *Emerging Waveguide Technology*; You, K.Y., Ed.; InTech: Houston, TX, USA, 2018; ISBN 978-1-78923-492-3.
23. Guzman-Sepulveda, J.R.; Castillo-Guzman, A.A. Wavelength Tuning of Multimode Interference Fiber Lasers: A Review. *Adv. Photonics Res.* **2021**, *2*, 2100051. [CrossRef]
24. Zhang, D.; Tan, Z. A Review of Optical Neural Networks. *Appl. Sci.* **2022**, *12*, 5338. [CrossRef]
25. Fu, T.; Zang, Y.; Huang, Y.; Du, Z.; Huang, H.; Hu, C.; Chen, M.; Yang, S.; Chen, H. Photonic Machine Learning with On-Chip Diffractive Optics. *Nat. Commun.* **2023**, *14*, 70. [CrossRef] [PubMed]
26. Meng, X.; Zhang, G.; Shi, N.; Li, G.; Azaña, J.; Capmany, J.; Yao, J.; Shen, Y.; Li, W.; Zhu, N.; et al. Compact Optical Convolution Processing Unit Based on Multimode Interference. *Nat. Commun.* **2023**, *14*, 3000. [CrossRef] [PubMed]
27. Ghosh, R.R.; Liu, D.; Reniers, S.; Jiao, Y.; Yao, W. Reconfigurable Optical Synaptic Weighting Engine Using a Liquid Crystal-Based Multimode Interference Coupler. *Front. Nanotechnol.* **2025**, *7*, 1593347. [CrossRef]
28. Yousefi, P.; Khalid, M.; Petruzzelli, V.; Calò, G. Design of Thin-Film Lithium Niobate Power Splitters and Combiners Based on Multimode Interference. *Opt. Quant. Electron.* **2025**, *57*, 198. [CrossRef]
29. Xu, M.; Chen, W.; He, M.; Wen, X.; Ruan, Z.; Chen, L.; Liu, L.; Yu, S.; Cai, X. Lithium Niobate Michelson Interferometer Modulator on Silicon-On-Insulator Platform. *arXiv* **2019**, arXiv:1905.01623.
30. Shen, H.; Fan, L.; Wang, J.; Wirth, J.C.; Qi, M. A Taper to Reduce the Straight-to-Bend Transition Loss in Compact Silicon Waveguides. *IEEE Photon. Technol. Lett.* **2010**, *22*, 1174–1176. [CrossRef]
31. Bai, B.; Yang, Q.; Shu, H.; Chang, L.; Yang, F.; Shen, B.; Tao, Z.; Wang, J.; Xu, S.; Xie, W.; et al. Microcomb-Based Integrated Photonic Processing Unit. *Nat. Commun.* **2023**, *14*, 66. [CrossRef] [PubMed]
32. Lin, Z.; Shastri, B.J.; Yu, S.; Song, J.; Zhu, Y.; Safarnejadian, A.; Cai, W.; Lin, Y.; Ke, W.; Hammood, M.; et al. 120 GOPS Photonic Tensor Core in Thin-Film Lithium Niobate for Inference and in Situ Training. *Nat. Commun.* **2024**, *15*, 9081. [CrossRef] [PubMed]
33. Guzmán-Sepúlveda, J.R.; Guzmán-Cabrera, R.; Castillo-Guzmán, A.A. Optical Sensing Using Fiber-Optic Multimode Interference Devices: A Review of Nonconventional Sensing Schemes. *Sensors* **2021**, *21*, 1862. [CrossRef] [PubMed]
34. Akil, D.S.; Othman, M.A.; Sherif, S.M.; Swillam, M.A. Designing an Optical Router Based on a Multimode-Interference Silicon-On-Insulator Coupler with Tunable Power Transmittance. *Photonics* **2024**, *11*, 221. [CrossRef]
35. Maese-Novo, A.; Halir, R.; Romero-García, S.; Pérez-Galacho, D.; Zavargo-Peche, L.; Ortega-Moñux, A.; Molina-Fernández, I.; Wangüemert-Pérez, J.G.; Cheben, P. Wavelength Independent Multimode Interference Coupler. *Opt. Express* **2013**, *21*, 7033. [CrossRef] [PubMed]
36. Yao, W.; Guo, F.; Zhong, M.; Lu, D. An Ultra-Compact InP 1310/1550 Nm Wavelength Division (De)Multiplexer Based on Channel-Shaped MMI Coupler. *Photonics* **2025**, *12*, 922. [CrossRef]
37. Gencal, H.; Aksoy, A.; Yigit, E.; Aydemir, U.; Demirtas, M. A Novel Deep Learning Approach for Accurate and Efficient Design of LNOI Power Splitters. *Opt. Quant. Electron.* **2026**, *58*, 239. [CrossRef]
38. Chen, G.; Ruan, Z.; Wang, Z.; Huang, P.; Guo, C.; Dai, D.; Chen, K.; Liu, L. Four-Channel CWDM Device on a Thin-Film Lithium Niobate Platform Using an Angled Multimode Interferometer Structure. *Photon. Res.* **2022**, *10*, 8. [CrossRef]
39. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 6999–7019. [CrossRef] [PubMed]
40. Van Willigenburg, L.G.; Stigter, J.D.; Molenaar, J. Sensitivity Matrices as Keys to Local Structural System Properties of Large-Scale Nonlinear Systems. *Nonlinear Dyn.* **2022**, *107*, 2599–2618. [CrossRef]
41. Mihalič, F.; Truntič, M.; Hren, A. Hardware-in-the-Loop Simulations: A Historical Overview of Engineering Challenges. *Electronics* **2022**, *11*, 2462. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.