

Supplementary Information

Multimode fiber laser cavities as nonlinear optical processors

Dilem Eşlik¹, Bahadır Utku Kesgin¹, Fatma Nur Kılınç²
Uğur Teğin^{1,2,*}

¹Department of Electrical and Electronics Engineering, Koç University, İstanbul, Türkiye

²Department of Computational Sciences and Engineering, Koç University, İstanbul, Türkiye

*Corresponding author: utegin@ku.edu.tr

Supplementary Note 1: Linear Discriminant Analysis (LDA) methodology

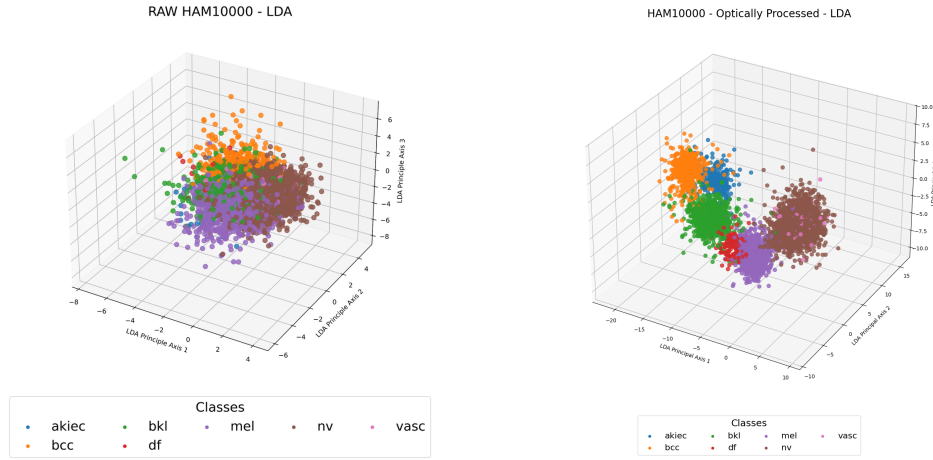
Linear Discriminant Analysis (LDA) is used as a visualization tool to assess the effect of optical preprocessing on class separability. LDA identifies linear projections that maximize the ratio of between-class variance to within-class variance and therefore provides an interpretable measure of linear separability in feature space.

LDA is applied to the pooled intensity features extracted from the steady-state output patterns of the multimode laser cavity. Prior to LDA, the output intensity images are spatially downsampled using non-overlapping average pooling with a fixed block size. The resulting reduced-resolution intensity maps are vectorized and normalized to ensure consistent scaling across samples.

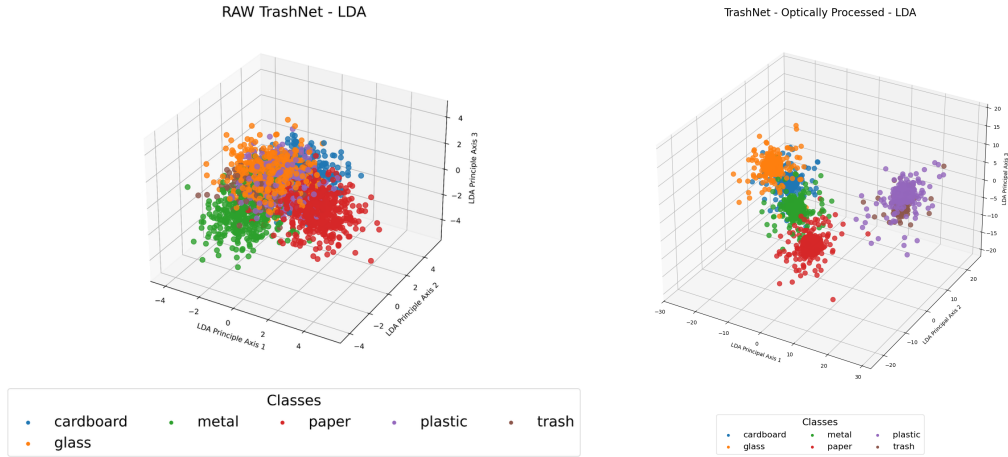
Class labels are encoded as integer values for analysis. LDA is performed using the singular value decomposition (SVD) solver, which does not require explicit estimation of the covariance matrix and is well suited for high-dimensional data. Up to three discriminant components are retained to enable three-dimensional visualization of the projected feature space.

The resulting LDA projections are used solely for qualitative visualization and are not involved in training or evaluating the classification models. As shown in Supplementary Figs. 1 and 2, LDA visualizations are presented for the HAM10000 and TrashNet datasets, respectively. In both cases, optically processed data exhibits compact and well-separated

clusters compared to raw digital inputs, indicating enhanced linear separability induced by the multimode laser cavity.



Supplementary Figure 1: Three-dimensional Linear Discriminant Analysis projections for the HAM10000 dataset before (left) and after (right) optical processing.



Supplementary Figure 2: Three-dimensional Linear Discriminant Analysis projections for the TrashNet dataset before (left) and after (right) optical processing.

Supplementary Note 2: Numerical implementation and convergence

Numerical simulations are performed using a (2+1)D beam propagation framework based on the split-step Fourier method (SSFM). The optical field is propagated through active and passive fiber sections while accounting for diffraction, refractive index contrast, intermodal coupling, and nonlinear gain saturation.

The computational window spans $54 \times 54 \mu\text{m}$ and is discretized on a uniform 128×128 grid. Propagation step sizes of $1 \mu\text{m}$ and $2 \mu\text{m}$ are used for the active Yb-doped fiber and the passive graded-index fiber, respectively. Absorbing boundary conditions are applied to suppress numerical reflections at the edges of the simulation window.

The two fiber sections are modelled with their physical transverse index profiles: a step-index profile for the Yb-doped gain fiber and a parabolic graded-index profile for the passive GRIN fiber. The cavity is simulated as an iterated round-trip map. Starting from the field encoded by the spatial light modulator (SLM), one round trip consists of: (i) propagation through the Yb-doped step-index section with the saturable-gain term of Eq. (1); (ii) propagation through the passive graded-index section; and (iii) the 50/50 coupler and Sagnac loop. The Sagnac loop is implemented by splitting the field at the coupler into the two counter-propagating loop arms, propagating each arm through the loop fiber, recombining them at the coupler with the appropriate $\pi/2$ coupler phase, and returning the reflected port into the cavity as the feedback field. The round trip is iterated, and convergence is declared when the relative change of the intracavity intensity profile between successive round trips falls below a fixed tolerance, which is reached in ≈ 100 round trips.

Cavity dynamics are modeled by iterating the propagation over multiple round trips. For each input phase pattern, the intracavity field converges to a stable steady-state intensity distribution. Convergence is verified by monitoring the relative change in intracavity intensity profiles over successive round trips.

Supplementary Note 3: Gain saturation, modal competition, and operating regime

The nonlinear response of the multimode laser cavity is dominated by gain saturation in the Yb-doped fiber. The local gain is modeled as an intensity-dependent function,

$$g = \frac{g_0}{1 + I/I_{\text{sat}}}, \quad (1)$$

where g_0 is the small-signal gain, I denotes the local optical intensity, and I_{sat} is the saturation intensity.

This gain saturation introduces mode competition, as spatial modes with higher localized intensity experience stronger gain depletion. Over successive cavity round trips, this mechanism redistributes power among modes in an input-dependent manner, resulting in distinct and reproducible steady-state intensity patterns.

The system operates in a continuous-wave regime at moderate intracavity intensities.

Under these conditions, instantaneous Kerr nonlinearities are negligible compared to gain-induced nonlinear effects, and the observed nonlinear feature transformation is governed primarily by gain saturation rather than ultrafast nonlinear phase modulation.

The laser-operation parameters are as follows. The saturation intensity of the Yb-doped fiber is $I_{\text{sat}} = h\nu/(\sigma\tau) \approx 9 \times 10^8 \text{ W m}^{-2}$ ($\lambda = 1030 \text{ nm}$, $\sigma \approx 2.5 \times 10^{-25} \text{ m}^2$, $\tau \approx 0.84 \text{ ms}$), equivalent to a saturation power of order 0.1–0.3 W over the $20 \mu\text{m}$ gain-fiber core. These intensities are well below the Kerr-relevant scale (Supplementary Note 10), consistent with attributing the nonlinear feature mapping to gain saturation.

Supplementary Note 4: SLM encoding and experimental data acquisition

Input data are encoded onto the intracavity optical field using a phase-only spatial light modulator operating as a programmable end mirror of the laser cavity.

Each image is embedded into a predefined region of the SLM active area, while the remaining pixels are held at zero phase. This fixed spatial placement ensures consistent illumination conditions and repeatable modal excitation across all samples. The pixel values of the embedded image are mapped linearly to 8-bit phase levels, forming a spatial phase mask.

Output intensity patterns are captured using a camera with fixed exposure time and gain settings to ensure consistent dynamic range across the dataset. Saturation checks are performed to exclude frames affected by detector clipping. Data acquisition is performed in batches and stored for subsequent digital processing and classification.

Supplementary Note 5: Feature extraction and dimensionality control

The steady-state output intensity patterns exhibit speckle-like spatial structure with high spatial redundancy. To reduce sensitivity to fine-scale fluctuations and control feature dimensionality, spatial average pooling is applied using non-overlapping blocks.

Prior to pooling, a fixed-size region of interest is cropped from each intensity image. Multiple crop locations are evaluated to identify regions with stable and informative spatial structure. For each crop, average pooling is performed with varying block sizes, resulting in reduced-resolution feature maps. The pooled intensity values are subsequently vectorized and normalized before classification.

Supplementary Note 6: Linear classifier and evaluation protocol

Classification is performed using a linear Ridge classifier as the electronic readout layer. The classifier operates on vectorized pooled intensity features and employs ℓ_2 regularization with a fixed regularization parameter.

For all datasets, the data are split into training and test sets using a stratified 80%–20% split to preserve class balance. Label encoding is applied to map class labels to integer indices. All features are normalized by their maximum value prior to training.

To ensure reproducibility, a fixed random seed is used for data splitting and classifier initialization. Performance is evaluated using classification accuracy, confusion matrices, and per-class precision and recall. Identical classifier architectures and evaluation protocols are used for both raw digital inputs and optically processed data.

Supplementary Note 7: Raw-image baseline preprocessing and classification

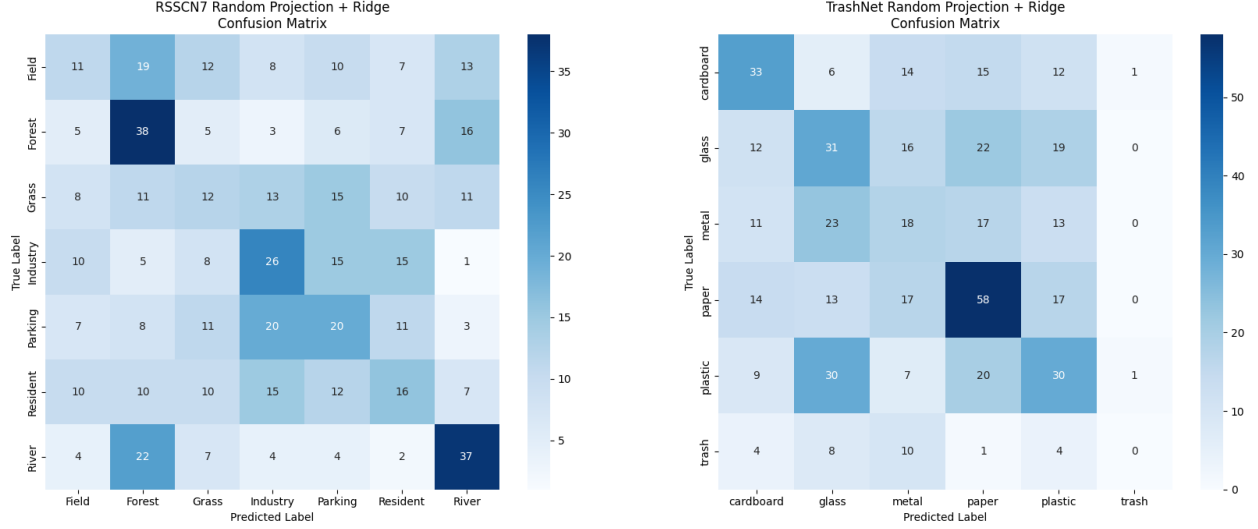
To provide a fair baseline comparison, classification is also performed directly on raw digital images without optical preprocessing. The same linear classifier architecture and evaluation protocol used for optically processed data are applied to the raw images.

All images are converted to grayscale and resized to a fixed spatial resolution of 600×600 pixels. No optical encoding, cavity processing, or phase modulation is applied in this baseline. The resized images are treated as intensity-only inputs.

To control feature dimensionality and ensure consistency with the optically processed pipeline, spatial average pooling is applied to the raw images using non-overlapping blocks. The pooled images are then vectorized to form feature vectors. Feature vectors are normalized prior to classification.

As an additional control experiment, Gaussian random projection baselines were evaluated for RSSCN7 and TrashNet. In these baselines, flattened raw-image feature vectors were projected onto lower-dimensional spaces using fixed Gaussian random matrices prior to Ridge classification. The projection dimensionality was chosen to approximately match the pooled feature dimensionality used in the optical pipeline. These controls test whether arbitrary high-dimensional random mixing alone is sufficient to reproduce the observed class separability. The resulting classification accuracies and confusion matrices remained substantially below those obtained using the multimode laser cavity, indicating that the cavity

generates structured nonlinear feature transformations beyond purely random projections.



Supplementary Figure 3: Confusion matrices for the Gaussian random-projection + Ridge baselines. Left: RSSCN7 dataset (28.57% accuracy). Right: TrashNet dataset (33.60% accuracy).

Supplementary Note 8: Dataset characteristics and image resolutions

Four benchmark datasets are used to evaluate the optical computing platform: TrashNet, RSSCN7, OCT MNIST, and HAM10000. These datasets span natural images, aerial scenes, and medical imaging, and differ in both image resolution and sample size.

TrashNet consists of 2,520 images of everyday waste objects with a native resolution of 224×224 pixels across six classes. RSSCN7 contains 2,400 aerial scene images with a resolution of 400×400 pixels distributed among seven scene categories. The HAM10000 dataset includes 10,015 dermoscopic skin lesion images with a resolution of 600×450 pixels across seven diagnostic classes. OCT MNIST comprises grayscale retinal OCT images with a native resolution of 28×28 pixels; a subset of 4,000 samples is used in this work to reduce computational overhead while preserving class balance. A summary of dataset characteristics, including class counts, sample sizes, and native image resolutions, is provided in Supplementary Table 3.

For experimental consistency, all images are converted to grayscale and resized to a uniform resolution of 600×600 pixels prior to encoding on the spatial light modulator. Pixel intensities are linearly mapped to 8-bit phase values to generate phase-only masks. The

same preprocessing pipeline is applied to both raw-image baselines and optically processed data to ensure fair comparison.

Supplementary Note 9: Modal gain competition, non-linear feature mapping, and processing capacity

S9.1 Modal description of the intracavity field. The intracavity optical field at any plane is expanded over the guided modes of the relevant fiber section,

$$A(x, y, z) = \sum_{m=1}^M a_m(z) \psi_m(x, y) e^{i\beta_m z},$$

where $\psi_m(x, y)$ are the transverse mode profiles, β_m the propagation constants, and $a_m(z)$ the complex modal amplitudes. The basis $\{\psi_m, \beta_m\}$ is that of the step-index Yb-doped fiber in the gain section and that of the graded-index multimode fiber in the passive section; at each splice the field is re-expanded in the modal basis of the following section. Free propagation over a length L advances each modal phase as $a_m \rightarrow a_m e^{i\beta_m L}$. Because the β_m differ between modes, this phase advance reshapes the transverse intensity through intermodal interference. This multimode-interference mechanism is an input-dependent but linear, energy-conserving redistribution of power among a fixed set of interference channels; on its own it implements a fixed linear (random-projection-like) transformation of the launched field.

S9.2 Gain saturation as a mode-coupling nonlinearity. The nonlinearity originates in the Yb-doped section, where the modal amplitudes are coupled through a shared, saturable inversion. The local gain depends on the total local intensity, $g(x, y) = g_0(x, y)/[1 + I(x, y)/I_{\text{sat}}]$ with $I(x, y) = |\sum_m a_m \psi_m|^2$. Projecting the saturated gain onto the modal basis, and treating the inversion in a lumped (mode-averaged) approximation, the gain experienced by mode m is, to leading order,

$$g_m = \frac{g_m^0}{1 + \frac{1}{P_{\text{sat}}} \sum_n \Gamma_{mn} P_n}, \quad \Gamma_{mn} = \int_{\text{core}} \psi_m(x, y) \psi_n(x, y) dx dy,$$

where $P_n = |a_n|^2$ and P_{sat} is the saturation power. The full numerical model retains the spatially resolved inversion; Eq. (S9.2) is the lumped modal reduction that makes the cross-saturation explicit. The off-diagonal terms Γ_{mn} ($m \neq n$) couple the gain of every mode to the power of every other mode: modes carrying more power deplete the shared inversion and suppress the gain available to competitors. This gain-mediated mode competition is the

differential, intensity-dependent nonlinearity referred to in the main text; it is intrinsically multimode and vanishes for a single mode, where it reduces to ordinary scalar gain saturation with no competition.

S9.3 Processing capacity and effective dimensionality. The number of independent optical features available to the readout is bounded by the modal content. A linear intensity readout of an M -mode field provides up to $\sim M$ independent intensity channels; pairwise intermodal beating contributes up to $\sim M(M+1)/2$ distinguishable second-order interference features, bounding the linearly decodable dimensionality of the speckle output. Gain competition makes the occupation of these features input-dependent and nonlinear, so the usable (class-separable) dimensionality grows with the number of participating modes up to this bound. After spatial average pooling, the readout operates on 2,500–10,000 features (main-text Table 1), well below the optical bound for a fiber supporting tens to hundreds of modes. The current capacity is therefore limited by the electronic readout and pooling stage, not by the optical feature space.

S9.4 Scaling with mode number. The number of guided modes scales with the normalised frequency $V = (2\pi a/\lambda) \text{NA}$ as $M \approx V^2/2$ (step index) and $M \approx V^2/4$ (graded index), i.e. $M \propto (a \text{NA}/\lambda)^2$, proportional to the core area and to NA^2 . Increasing the core radius or numerical aperture enlarges the feature space, but at two costs: the number of round trips required to reach steady state increases, and the sensitivity of the steady-state pattern to environmental perturbation grows. This defines a capacity–stability trade-off that bounds the useful mode count for a given calibration interval.

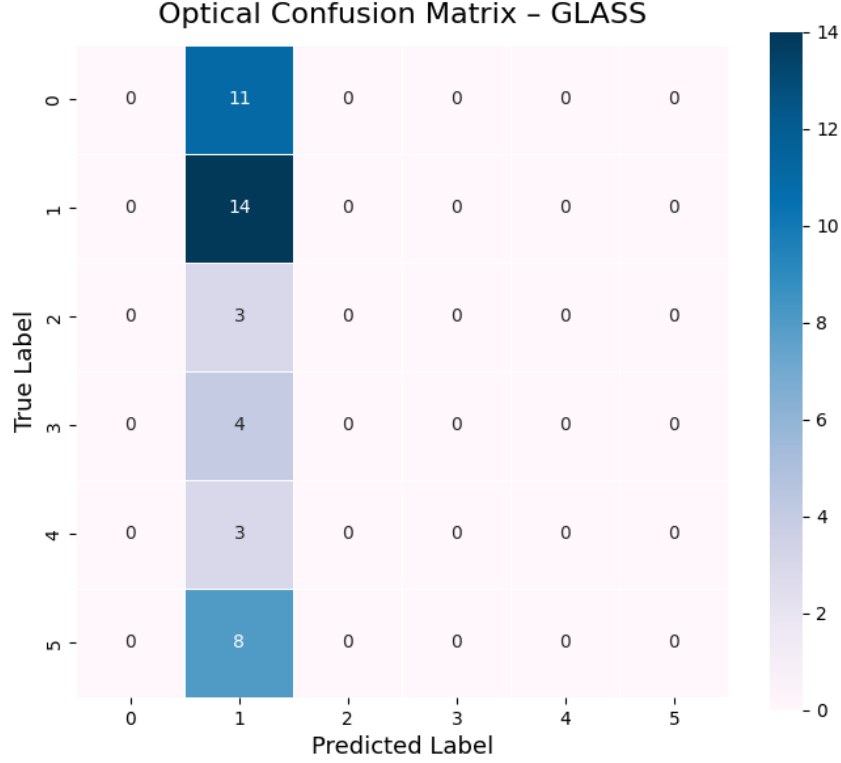
S9.5 Number of modes participating in the dynamics. The passive graded-index section supports on the order of hundreds of guided modes, whereas the step-index Yb-doped section supports only a few transverse modes. Because the graded-index fiber is spliced on-axis to the few-mode output of the gain section, only the fundamental and low-order higher-order modes of the GRIN fiber are appreciably excited, and the high-order GRIN modes receive negligible power. The set of modes that participates in the cavity dynamics is thus bounded above by the GRIN support and below by the few-mode Yb support, and is further shaped by the SLM phase pattern, which diversifies the launched mode content from round trip to round trip. The number of participating modes is therefore of order ten to a few tens—bounded above by the GRIN support and below by the few transverse modes of the Yb gain fiber. The high-dimensional feature space exploited for classification arises from the interference of this participating set together with the input-dependent gain competition among its members.

Supplementary Note 10: Mechanism controls — Kerr-phase estimate, gain-off comparison, and spectral evolution

S10.1 Estimate of the Kerr phase shift. The instantaneous Kerr nonlinearity produces a single-pass nonlinear phase $\Delta\phi_{\text{Kerr}} = (2\pi/\lambda) n_2 I L$, where $n_2 \approx 2.6 \times 10^{-20} \text{ m}^2 \text{ W}^{-1}$ is the nonlinear index of silica, I the intracavity intensity, and L the interaction length. For an intracavity power of order 0.1 W over the $\sim 50 \mu\text{m}$ -core graded-index fiber (effective area $\sim 2 \times 10^{-9} \text{ m}^2$), $I \sim 5 \times 10^7 \text{ W m}^{-2}$; over an interaction length $L \sim 5 \text{ m}$ this gives $\Delta\phi_{\text{Kerr}} \sim 4 \times 10^{-5} \text{ rad}$ per pass, and of order 10^{-3} rad accumulated over the ~ 100 round trips required to reach steady state. This is four to five orders of magnitude below the π -scale phase modulation imposed by the SLM and below the phase redistribution produced by gain competition, which justifies neglecting the Kerr term in Main Text Equation (3).

S10.2 Why a pump-off control is not applicable. The pump is free-space coupled into the cavity through a dichroic mirror, and the resonator is formed between the SLM and the free-space output path. The pumped Yb-doped fiber is the only optical source in the system, and there is no externally injected seed. Switching the pump off, or replacing the gain fiber with a passive fiber, therefore removes all intracavity light, leaving no field to characterise. The role of gain is instead established by the numerical control described below, in which a seed field is propagated through the cavity with the gain term disabled.

S10.3 Gain-off numerical control. To isolate the contribution of gain competition from that of linear multimode interference, the digital twin propagates the SLM-encoded field through the cavity with the saturable-gain operator disabled, retaining only diffraction and intermodal mixing. Without gain, the steady-state output reduces to a fixed linear mixing of the launched field and the class-separable structure collapses under the same round-trip conditions, which is verified by the simulations. This confirms that the input-dependent feature separation is produced by gain-mediated mode competition rather than by linear multimode interference alone.



Supplementary Figure 4: Confusion matrix obtained when the saturable-gain term is disabled ($g = 0$), leaving only linear multimode propagation and diffraction. The resulting classification accuracy decreases to approximately 32.56%, indicating that linear multimode mixing alone is insufficient to generate the nonlinear feature separation observed in the active cavity.

Supplementary Note 11: Energy and throughput analysis

S11.1 Power budget and energy per inference. The component-level electrical power budget of the proof-of-concept setup is given in Supplementary Table 1. The dominant contribution is the liquid-crystal spatial light modulator (SLM) driver; the pump diode and the host/camera overhead are minor. The current setup consumes ≈ 32.5 W in total. At the SLM-limited input frame rate of 60 Hz, the energy per inference is $32.5 \text{ W}/60 \text{ Hz} \approx 0.54 \text{ J/inference}$.

Supplementary Table 1: Component-level electrical power budget for the multimode-fiber-laser proof-of-concept setup.

Component	Power (W)
Pump laser diode (976 nm, 1.2 W optical)	5 (electrical)
Phase-only SLM (Holoeye Pluto 2.1)	24
Host computer and camera (FLIR BFS-U3-04S2M-CS)	3.5
System total	≈ 32.5

S11.2 Throughput and efficiency. At the current 60 Hz operating point, the throughput is $R_o \times 60 \approx 1.29 \times 10^{13}$ OPS/s = 12.9 TOPS, i.e. $12.9/32.5 \approx 0.40$ TOPS/W (Supplementary Table 2). The gap to digital accelerators is dominated by the SLM-limited modulation rate, not by the optical computation. The cavity round-trip time is $t_{rt} = nL/c \approx 50$ ns for the ~ 10 m round-trip optical path, so the ~ 100 -round-trip buildup is $\sim 5 \mu\text{s}$, setting an upper bound of $\sim 2 \times 10^5$ inferences per second. At a representative 32 kHz modulation rate and ≈ 30 W, the throughput is $\approx 6.9 \times 10^3$ TOPS (6.9 POPS) at ≈ 230 TOPS/W, exceeding the NVIDIA A100 (1.56 TOPS/W) by more than two orders of magnitude.

Supplementary Table 2: Throughput and energy efficiency of the proposed optical processor and a digital baseline. The projected row assumes that the modulator no longer limits the operating rate; the cavity buildup time sets the ultimate ceiling at $\sim 2 \times 10^5$ inferences per second.

Scenario	P (W)	Rate (Hz)	Throughput	Efficiency (TOPS/W)
This work — SLM-based (current)	32.5	60	12.9 TOPS	0.40
This work — fast modulator (projected)	≈ 30	3.2×10^4	6.9×10^3 TOPS	≈ 230
NVIDIA A100 (INT16)	—	—	312 TOPS	1.56

Supplementary Tables

Supplementary Table 3: Datasets used in the experimental demonstrations and their native image resolutions.

Dataset	Classes	Sample size	Image resolution	Domain
TrashNet	6	2,520	224×224	Waste classification
RSSCN7	7	2,400	400×400	Aerial scene recognition
OCT MNIST	4	4,000	28×28	Retinal imaging
HAM10000	7	10,015	600×450	Dermoscopic lesions

Supplementary Table 4: Comparison of trainable complexity and classification accuracy on RSSCN7 under a 50%–50% train–test split. CNN baselines are taken from Sun et al. [1]. The raw-pixel and Gaussian random-projection baselines are digital controls of complexity comparable to the proposed multimode laser (MML) readout and are evaluated under the identical classification pipeline.

Learner	Trainable parameters	RSSCN7 accuracy
MML computing (this work)	2.5×10^3 – 1.0×10^4	94.21%
Raw-pixel + Ridge (linear)	$\sim 10^3$ – 10^4	18.75%
Gaussian random projection + Ridge	$\sim 10^3$ – 10^4	28.57%
DCNN [1]	deep ($\gtrsim 10^7$)	77.00%
TDFF [1]	deep ($\gtrsim 10^7$)	92.37%
DAC [1]	deep ($\gtrsim 10^7$)	93.43%
AlexNet [1]	6.0×10^7	91.85%
VGG16 [1]	1.38×10^8	93.57%
ResNet-50 [1]	2.6×10^7	93.64%
TEX-Net-LF [1]	deep ($\gtrsim 10^7$)	94.00%
GLNet (VGG16 backbone) [1]	$> 1.38 \times 10^8$	95.07%

Supplementary Table 5: Per-class precision, recall, and F1 for optically processed HAM10000 (80/20 stratified split; 2,003 test images), computed from the confusion matrix in Fig. 4d.

Class	Precision	Recall	F1	Support (test)
akiec	0.8312	0.9846	0.9014	65
bcc	0.8491	0.8738	0.8612	103
bkl	0.8545	0.8273	0.8406	220
df	0.5405	0.8696	0.6667	23
mel	0.5920	0.8655	0.7031	223
nv	0.9982	0.8449	0.9152	1341
vasc	0.2385	0.9286	0.3796	28
Overall accuracy	0.8527 (1708/2003)			
Macro precision	0.7006			
Balanced accuracy (macro recall)	0.8849			
Macro F1	0.7525			

Supplementary Table 6: Per-class precision, recall, and F1 for optically processed OCT-MNIST (800 test images; 200 samples per class), computed from the confusion matrix in Fig. 4b.

Class	Precision	Recall	F1	Support (test)
CNV	1.0000	1.0000	1.0000	200
DME	1.0000	0.9650	0.9822	200
Drusen	0.9569	1.0000	0.9780	200
Normal	1.0000	0.9900	0.9950	200
Overall accuracy	0.9888 (791/800)			
Macro precision	0.9892			
Balanced accuracy (macro recall)	0.9888			
Macro F1	0.9888			

Supplementary References

- [1] H. Sun, Y. Lin, Q. Zou, S. Song, J. Fang, and H. Yu, “Convolutional neural networks based remote sensing scene classification under clear and cloudy environments,” in *Proc. IEEE/CVF Int. Conf. Computer Vision Workshops (ICCVW)* (IEEE, 2021), pp. 713–720.