

Multimode fiber laser cavities as nonlinear optical processors

Corresponding Author: Professor Ugur Tegin

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This manuscript presents an experimental demonstration of a multimode fiber laser cavity used as a nonlinear optical processor for image classification. The approach uses multimode interference and gain saturation to map input phase patterns into spatial intensity distributions, followed by a linear classifier.

While the work is technically interesting and the experimental results are also interesting, the overall contribution remains incremental and does not reach the level of novelty or physical advances expected for Communications Physics.

The main idea closely follows established areas of research in optical computing and multimode fiber processing. The use of nonlinear cavity dynamics as a feature mapping mechanism is well known, and the manuscript does not introduce a new concept or mechanism beyond this existing body of work. The work is an interesting implementation.

In addition, the manuscript provides limited physical description into the underlying processes. The role of multimode interference and gain saturation is described, but there is no deeper analysis of modal dynamics, scaling behavior or processing capacity. As a result, the work does not significantly advance our understanding of the physics governing these systems.

The reported performance is based on relatively simple datasets, and comparisons to neural networks are not carried out under equivalent conditions. This makes it difficult to understand the true significance of the results. Moreover, system advantages such as energy efficiency, latency, and scalability are not presented, which weakens the broader claims regarding practical optical computing.

Overall, the manuscript presents a very good experimental study, but the level of novelty and impact is more appropriate for a specialized optics journal rather than a high-impact physics journal.

Reviewer #2

(Remarks to the Author)

There is significant interest in the study of optical analogue computers exploiting the parallel processing capability of optics, with the aim of reducing the energy requirements of conventional digital processors, particularly for complex tasks. This manuscript presents an original and promising approach, based on a fiber laser cavity composed by multimode active and passive fibers. Spatial information is encoded by a means of a spatial light modulator (SLM) which provides a programmable mirror in the laser cavity. Different images encode phase patterns on the SLM produce different intensity distributions at the laser output, which are stable and well separable, demonstrating the capability of the device for image classification with relatively good accuracy. The main advantage of the implementation appears to be the strong reduction in the number of trainable parameters with respect to traditional neural network approaches. On the other hand, the processing speed is limited in this implementation by the low refresh rate of the liquid crystal based SLM. Overall, the demonstration is novel and convincing, and it represents a useful advance with respect to the state of the art to deserve publication. That said, I have some remarks on the numerical simulation of the laser cavity and on the operation of the fiber laser that would deserve additional clarifications by the Authors:

1) Eq.(1) is said to include diffraction and intensity dependent gain saturation, what about the refractive index profile of the two fibers that compose the cavity? Besides, it is not clear how the Sagnac loop was numerically implemented, did the Authors simulate a map of the main cavity and the loop until convergence was obtained? Simulation details and illustrative graphics should be added to the supplementary note S2.

2) It is claimed that nonlinearity is introduced by mode-differential gain saturation. Was that verified either experimentally and numerically? What is the saturation intensity? What is the intracavity intensity and the output power from the laser? In section 4.2 is written that experimental details are contained in note S6 but this note does not provide experimental details about the laser operation. What was the threshold power of the laser?

3) What kind of 50/50 coupler was used between the active fiber and the GRIN fiber? Is it a multimode fiber coupler? What are the input and output fibers, graded index or step index, and what are their core/cladding sizes?

4) The Yb fiber supports a few transverse modes, unlike the GRIN fiber that supports hundreds of modes. If the GRIN fiber was spliced on-axis with the output fibers from the coupler, then only the fundamental mode and a few HOM of the GRIN fiber will be excited. Can the Authors estimate the number of spatial modes involved in the cavity dynamics?

5) In CW operation, the cavity operates on a single spatial mode, or it lases simultaneously on many spatial modes at different frequencies?

Reviewer #3

(Remarks to the Author)

The manuscript presents a continuous-wave Yb-doped multimode fiber laser cavity as a physical nonlinear feature extractor for image classification. A spatial light modulator serves as a programmable cavity end mirror, the cavity is closed by a Sagnac loop, and the steady-state output intensity is read out with a linear Ridge classifier. The idea is interesting and the experimental effort is clearly substantial, but several issues require major revision.

1. The main mechanistic claim is not yet convincingly established. The manuscript attributes the nonlinear mapping to gain saturation and multimode interference while neglecting Kerr effects, but the key controls are missing. The authors should show results with the pump turned off and, ideally, with the Yb-doped section replaced by a passive fiber. A simple estimate of the Kerr phase shift, together with output spectra as a function of pump power, would also help support the physical interpretation.

2. The baseline comparison is too weak. A linear classifier on raw pixels is not an adequate reference for a random-feature optical system. Stronger baselines are needed, in particular a matched random-projection-plus-Ridge baseline and a small digital model with comparable trainable complexity.

3. The OCT MNIST result needs clarification. The reported accuracy is unusually high, and the use of a 4,000-image subset is not clearly explained. The authors should state explicitly which split was used, whether the reported metric is accuracy or AUC, and ideally verify the result using the standard MedMNIST evaluation pipeline.

4. The statistical reporting is insufficient. All results appear to be single-point estimates. The authors should report performance over multiple random splits or seeds, and ideally across repeated experimental sessions, given the likely sensitivity to drift and calibration.

5. The LDA analysis in Fig. 5 may not be valid if the projection was fit using both training and test data. LDA should be fit on the training set only, with held-out test samples projected afterward.

6. The HAM10000 result should be reported with class-balanced metrics. Since this dataset is imbalanced, overall accuracy alone is not sufficient. Balanced accuracy and per-class F1 scores should also be provided.

7. The gap between the simulation and experimental improvements is not discussed. In the UCI Glass simulation, the gain is modest, whereas in the experimental image datasets the reported gains over the raw baselines are much larger. The authors should comment on why the improvement differs so strongly across these cases.

8. Important recent literature on passive recurrent optical nonlinear processing is missing and should be discussed, especially because it bears directly on the manuscript's framing.

9. The parameter-count comparison in Table 1 is not fully fair. The manuscript compares readout parameters against the full parameter counts of CNN models. For this type of claim, energy per inference and throughput would be more informative metrics.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The Authors did not provide a point to point response letter to the previous reviewers' reports. Therefore, it is not possible to assess if the revision was made according to the previous reports.

Reviewer #3

(Remarks to the Author)

The revised manuscript has addressed my previous concerns satisfactorily.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Reviewers
Multimode fiber laser cavities as nonlinear optical processors

We would like to thank all three Reviewers for carefully evaluating our work and for the constructive comments that have improved the quality and presentation of our study. We have responded to every comment and question in detail below. In this document the listed colours indicate the following:

- **Blue:** Comment of the Reviewer.
 - **Black:** Response of the Authors.
 - **Red:** Changes in the Manuscript.
 - **Purple:** Changes in the Supplementary Material.
-

Reviewer 1

This manuscript presents an experimental demonstration of a multimode fiber laser cavity used as a nonlinear optical processor for image classification. The approach uses multimode interference and gain saturation to map input phase patterns into spatial intensity distributions, followed by a linear classifier. While the work is technically interesting and the experimental results are also interesting, the overall contribution remains incremental and does not reach the level of novelty or physical advances expected for Communications Physics.

The main idea closely follows established areas of research in optical computing and multimode fiber processing. The use of nonlinear cavity dynamics as a feature mapping mechanism is well known, and the manuscript does not introduce a new concept or mechanism beyond this existing body of work. The work is an interesting implementation.

In addition, the manuscript provides limited physical description into the underlying processes. The role of multimode interference and gain saturation is described, but there is no deeper analysis of modal dynamics, scaling behavior or processing capacity. As a result, the work does not significantly advance our understanding of the physics governing these systems.

The reported performance is based on relatively simple datasets, and comparisons to neural networks are not carried out under equivalent conditions. This makes it difficult to understand the true significance of the results. Moreover, system advantages such as energy efficiency, latency, and scalability are not presented, which weakens the broader claims regarding practical optical computing.

Overall, the manuscript presents a very good experimental study, but the level of novelty and impact is more appropriate for a specialized optics journal rather than a high-impact physics journal.

We thank the Reviewer for the careful reading and for recognising the technical and experimental merit of the work. The Reviewer raises four connected concerns, which we address as Comments 1–4 below: (1) the conceptual novelty of the platform; (2) the absence of a deeper physical analysis of modal dynamics, scaling, and processing capacity; (3) the comparison to neural networks under equivalent conditions; and (4) the lack of system-level figures of merit. We have added a new theoretical treatment of gain-saturated modal competition (Supplementary Note S9) and a quantitative energy/throughput analysis (Supplementary Note S11) that directly target these points.

Reviewer 1, Comment 1 — Conceptual novelty

The use of nonlinear cavity dynamics as a feature mapping mechanism is well known, and the manuscript does not introduce a new concept or mechanism beyond this existing body of work.

We respectfully clarify the specific advance of this work, which we agree was under-stated in the original framing. Existing approaches to nonlinearity in multimode-optical computing fall into three families. (i) *Passive* speckle/scattering systems realise a fixed linear random projection of the launched field [8]. (ii) *Pulsed* nonlinear multimode systems obtain their nonlinearity from Kerr-mediated spatiotemporal dynamics at high peak intensity [9, 10]. (iii) *Structural* approaches synthesise an effective nonlinearity from linear elements arranged in a cavity or multi-plane geometry, for example multi-plane light conversion and recurrent linear scattering [6, 7]. Our platform is distinct from all three: it uses an *active*, continuous-wave laser cavity in which the nonlinearity is mode-competition-driven cross-gain saturation, and in which intracavity feedback makes the input-to-feature map *recurrent* rather than single-pass. The computational primitive is therefore a continuous-wave, low-power, gain-saturated nonlinear fixed-point map. We have sharpened the Introduction to state this distinction explicitly, supported by the new theory in Supplementary Note S9 and by the random-projection baselines (Reviewer 3, Comment 2), which confirm empirically that a fixed linear projection cannot reproduce the measured performance. The distinction also extends to our group’s own recent optical processors: unlike a passive, genetically-programmed optical random network [3] or pulsed multimode dynamics operated near the edge of spatiotemporal chaos [2], the present system is an *active, continuous-wave, gain-saturated* cavity operated in a stable, deterministic fixed-point regime. The concrete capability this enables, and which the passive and pulsed approaches do not, is a reproducible, low-power nonlinear feature map whose nonlinearity is supplied by the gain medium itself, so that a single fixed linear readout decodes stable class-separable features without any external nonlinearity, high-peak-power pulses, or trained optical elements.

Added to the Introduction:

Also, structural nonlinearities from linear optical elements arranged in a recurrent or multi-plane geometry, can be harnessed as a computing media [13, 12, 7, 6, 21]. In these optical processors the nonlinearity is either supplied by an external digital stage or emerges structurally from the recurrence of linear optical elements.

Unlike previously reported, the optical feature extraction in this system occurs from gain-saturated fixed-point operation realised at low average power.

Reviewer 1, Comment 2 — Modal dynamics, scaling, and processing capacity

The role of multimode interference and gain saturation is described, but there is no deeper analysis of modal dynamics, scaling behavior or processing capacity.

We agree, and this comment prompted the most substantial addition to the revised manuscript: a dedicated theoretical treatment of the modal dynamics in Supplementary Note S9, and a corresponding paragraph in the main-text Discussion. The note (i) writes the cavity field in a modal basis, (ii) derives the cross-gain saturation that couples the modes and shows why it is a genuine, intrinsically multimode nonlinearity, (iii) frames the steady state as the fixed point of a recurrent nonlinear map, (iv) bounds the processing capacity and effective dimensionality, and (v) gives the scaling with mode number, including an estimate of the number of modes participating in the dynamics. The full text added to the manuscript and the Supplementary Material follows.

Added to the Methods:

The computational nonlinearity of the cavity originates in mode-differential gain saturation. The intracavity field is a superposition of guided modes,

$$A(x, y, z) = \sum_m a_m(z) \psi_m(x, y) e^{i\beta_m z} \quad (1)$$

whose phases advance at different rates so that multimode interference redistributes the transverse intensity in an input-dependent but linear, energy-conserving manner. The nonlinearity enters through the Yb-doped section, where all modes draw on a common, saturable inversion: the gain seen by mode m is expressed as

$$g_m = g_m^0 / (1 + P_{\text{sat}}^{-1} \sum_n \Gamma_{mn} P_n) \quad (2)$$

with modal power P_n and the overlap of modes m and n (Γ_{mn}) within the doped core. The off-diagonal overlaps couple the gain of every mode to the power of every other mode, so that a mode carrying more power suppresses the gain of its competitors. Cavity feedback iterates this nonlinear, mode-competitive map to an input-dependent steady state. The dimensionality of this representation is bounded by the modal content, up to $\sim M$ intensity channels and $\sim M(M+1)/2$ second-order interference features for M participating modes, and scales as $M \propto (a \text{ NA}/\lambda)^2$. A complete derivation, together with an estimate of the number of participating modes, is given in Supplementary Note 9.

New Supplementary Note S9 (full text):

Supplementary Note S9: Modal gain competition, nonlinear 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 of order 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.

Reviewer 1, Comment 3 — Comparison to neural networks under equivalent conditions

The reported performance is based on relatively simple datasets, and comparisons to neural networks are not carried out under equivalent conditions.

We agree that the original comparison mixed evaluation protocols, and we have strengthened it in three ways. First, we add a matched Gaussian random-projection + Ridge baseline, evaluated with the same readout and split as the optical features (Reviewer 3, Comment 2): it reaches only 28.57% (RSSCN7) and 33.60% (TrashNet), confirming that a fixed linear projection of comparable output dimensionality cannot account for the performance. Second, we report the full set of published CNN baselines for RSSCN7 under the same 50%–50% split used for the multimode-laser (MML) system (new Supplementary Table S4), drawn from the benchmark study of Sun et al. [16]: AlexNet 91.85%, VGG16 93.57%, ResNet-50 93.64%, DCNN 77.00%, DAC 93.43%, TDFF 92.37%, TEX-Net-LF 94.00%, and GLNet variants 93.78–95.07%, all using 1.5×10^7 – 10^8 trainable parameters, against 94.21% for the MML system with 2,500–10,000 readout parameters. Third, we note that the platform belongs to the same family as other optical neural networks that offload feature extraction to optics with a compact electronic readout—diffractive deep neural networks [5], fiber-based diffractive networks [4], and multi-plane structural processors [6]—and that the MML reaches competitive accuracy with a trainable-parameter budget at the low end of this group.

We also note, for balance, that the comparison does not favour the MML system on every

dataset. On the imbalanced, fine-grained HAM10000 task the deep CNNs reach higher accuracy (VGG16 96.05%, AlexNet 95.29%) than the MML system (85.22%, balanced accuracy 0.701), and we do not claim accuracy parity there. The defensible claim on that task is parameter and energy efficiency rather than absolute accuracy: the MML attains 85% with a 10^3 – 10^4 -parameter linear readout against 10^7 – 10^8 parameters for the CNNs. We have made this asymmetry explicit in the revised Discussion so that the cross-architecture comparison is stated even-handedly.

New Supplementary Table S4 (RSSCN7 model comparison, 50%–50% split):

Learner	Trainable parameters	RSSCN7 accuracy
MML computing (this work)	2,500–10,000	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 [Sun et al.]	deep ($\gtrsim 10^7$)	77.00%
TDFF [Sun et al.]	deep ($\gtrsim 10^7$)	92.37%
DAC [Sun et al.]	deep ($\gtrsim 10^7$)	93.43%
AlexNet [Sun et al.]	6.0×10^7	91.85%
VGG16 [Sun et al.]	1.38×10^8	93.57%
ResNet-50 [Sun et al.]	2.6×10^7	93.64%
TEX-Net-LF [Sun et al.]	deep ($\gtrsim 10^7$)	94.00%
GLNet (VGG16 backbone) [Sun et al.]	$> 1.38 \times 10^8$	95.07%

Supplementary Table S4: Comparison of trainable complexity and accuracy on RSSCN7 under a 50%–50% train–test split. CNN baselines are taken from Sun et al. [16]. The raw-pixel and random-projection rows are digital baselines of complexity comparable to the MML readout, evaluated under the identical pipeline.

Reviewer 1, Comment 4 — Energy efficiency, latency, and scalability

Moreover, system advantages such as energy efficiency, latency, and scalability are not presented, which weakens the broader claims regarding practical optical computing.

We have added a quantitative system-level analysis (new Supplementary Note S11) that reports a component-level power budget, the energy per inference and throughput at the current operating point, and the projected figures under a faster modulator, together with a comparison to a digital baseline. This also answers Reviewer 3, Comment 9. The analysis adopts the operations-counting framework of Xu et al. [19] used for diffractive optical processors so that the optical and digital figures are directly comparable. The full text added to the manuscript and the Supplementary Material follows.

Added to the Discussion:

At the current operating point the system consumes approximately 32.5 W in total (pump diode ≈ 5 W; SLM driver ≈ 24 W; host computer and camera ≈ 3.5 W). At the SLM-limited input frame rate of 60 Hz this corresponds to ≈ 0.54 J per inference and a throughput of ≈ 12.9 TOPS, or ≈ 0.40 TOPS/W (Supplementary Note S11). This figure is bottlenecked by the liquid-crystal SLM refresh rate and the digital host, not by the optical computation: the cavity performs the high-dimensional nonlinear feature transformation intrinsically during optical propagation and gain evolution, and the marginal optical cost per inference is negligible.

New Supplementary Note S11 (full text):

Supplementary Note S11: 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 S5. The dominant contribution is the liquid-crystal 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 $32.5 \text{ W}/60 \text{ Hz} \approx 0.54 \text{ J/inference}$.

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

Supplementary Table S5: Component-level electrical power budget for the multimode-fiber-laser proof-of-concept setup.

S11.2 Throughput and efficiency. At the current 60 Hz operating point, the throughput is $R_o \times 60 \approx 1.29 \times 10^{13} \text{ OPs/s} = 12.9 \text{ TOPS}$, i.e. $12.9/32.5 \approx 0.40 \text{ TOPS/W}$ (Supplementary Table S6). The gap to digital accelerators is dominated by the SLM-limited modulation rate, not by the optical computation. The cavity round-trip time is $t_{\text{rt}} = nL/c \approx 50 \text{ ns}$ for the $\sim 10 \text{ m}$ 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 $\approx 30 \text{ W}$, the throughput is $\approx 6.9 \times 10^3 \text{ TOPS}$ (6.9 POPS) at $\approx 230 \text{ TOPS/W}$, exceeding the NVIDIA A100 (1.56 TOPS/W) by more than two orders of magnitude.

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	$6.9 \times 10^3 \text{ TOPS}$	≈ 230
NVIDIA A100 (INT16)	—	—	312 TOPS	1.56

Supplementary Table S6: Throughput and energy efficiency of this work and a digital baseline. The projected row assumes the modulator no longer limits the rate; the cavity buildup time sets the ultimate ceiling at $\sim 2 \times 10^5$ inferences per second.

Reviewer 2

There is significant interest in the study of optical analogue computers exploiting the parallel processing capability of optics, with the aim of reducing the energy requirements of conventional digital electronic processors, particularly for complex machine learning tasks. This manuscript presents an original and promising approach, based on a fiber laser cavity composed by multimode active and passive fibers. Spatial information is encoded by a means of a spatial light modulator (SLM) which provides a programmable mirror in the laser cavity. Different images encodes a phase patterns on the SLM produce different intensity distributions at the laser output, which are stable and well separable, demonstrating the capability of the device for image classification with relatively good accuracy. The main advantage of the implementation appears to be the strong reduction in the number of trainable parameters with respect to traditional neural network approaches. On the other hand, the processing speed is limited in this implementation by the low refresh rate of the liquid crystal based SLM. Overall, the demonstration is novel and convincing, and it represents a useful advance with respect to the state of the art to deserve publication. That said, I have some remarks on the numerical simulation of the laser cavity and on the operation of the fiber laser that would deserve additional clarifications by the Authors:

We thank the Reviewer for the positive overall assessment and for the detailed, technically precise questions, which we address point by point below.

Reviewer 2, Comment 1 — Index profiles and Sagnac-loop implementation

Eq.(1) is said to include diffraction and intensity dependent gain saturation, what about the refractive index profile of the two fibers that compose the cavity? Besides, it is not clear how the Sagnac loop was numerically implemented, did the Authors simulate a map of the main cavity and the loop until convergence was obtained? Simulation details and illustrative graphics should be added to the supplementary note S2.

We thank the Reviewer for this important point. The two fibers are modelled with their physical index profiles: a step-index profile for the Yb-doped gain fiber and a parabolic graded-index profile for the passive GRIN fiber. The transverse index distribution enters the diffraction operator $\hat{D}$ through the guiding term, so that the modal basis used in each section is that of the respective fiber rather than free space. The cavity is simulated as an iterated round-trip map, not a single pass: each round trip applies SLM phase encoding, propagation through the Yb-doped section with saturable gain, propagation through the GRIN section, and the 50/50 coupler/Sagnac loop, and the map is iterated until the relative change of the intracavity intensity profile between successive round trips falls below a fixed tolerance. We have added these details to Note S2.

Added to Supplementary Note S2:

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 SLM-encoded field, 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.

Reviewer 2, Comment 2 — Verification of the gain-saturation nonlinearity and laser operating parameters

It is claimed that nonlinearity is introduced by mode-differential gain saturation. Was that verified either experimentally and numerically? What is the saturation intensity? What is the intracavity intensity and the output power from the laser? In section 4.2 is written that experimental details are contained in note S6 but this note does not provide experimental details about the laser operation. What was the threshold power of the laser?

We thank the Reviewer for these precise questions. The saturation intensity of the Yb-doped fiber is $I_{\text{sat}} = h\nu/(\sigma\tau)$, with photon energy $h\nu$, effective emission cross-section σ , and upper-state lifetime τ ; using $\lambda = 1030$ nm, $\sigma \approx 2.5 \times 10^{-25}$ m² and $\tau \approx 0.84$ ms for Yb:silica [15], $I_{\text{sat}} \approx 9 \times 10^8$ W m⁻², corresponding to P_{sat} of order 0.1–0.3 W over a 20 μ m core. This places the operating point in the saturated-gain regime while remaining far below any Kerr-relevant intensity (Reviewer 3, Comment 1). The mode-differential nature of the saturation is verified numerically by the gain-off control of Supplementary Note S10: disabling the gain term collapses the feature separation toward the random-projection baseline. We have added the requested laser-operation parameters to Note S6.

Added to Supplementary Note S3:

The laser-operation parameters are as follows. The saturation intensity of the Yb-doped fiber is $I_{\text{sat}} = h\nu/(\sigma\tau) \approx 9 \times 10^8$ W m⁻² ($\lambda = 1030$ nm, $\sigma \approx 2.5 \times 10^{-25}$ m², $\tau \approx 0.84$ ms), equivalent to a saturation power of order 0.1–0.3 W over the 20 μ m gain-fiber core. These intensities are well below the Kerr-relevant scale (Supplementary Note S10), consistent with attributing the nonlinear feature mapping to gain saturation.

Reviewer 2, Comment 3 — Coupler and fiber specifications

What kind of 50/50 coupler was used between the active fiber and the GRIN fiber? Is it a multimode fiber coupler? What are the input and output fibers, graded index or step index, and what are their core/cladding sizes?

We thank the Reviewer and have added these specifications to the Methods. The 50/50 coupler is a multimode fiber coupler with graded-index multimode pigtails matched to the cavity GRIN fiber. The gain section is a step-index Yb-doped fiber (20/125 μm core/cladding); the passive section is a graded-index multimode fiber (50/125 μm core/cladding).

Added to Methods (Experimental setup):

The cavity uses a 50/50 multimode fiber coupler with graded-index multimode pigtails (50/125 μm) matched to the passive cavity fiber. The gain section is a step-index Yb-doped fiber (20/125 μm core/cladding); the passive section is a graded-index multimode fiber (50/125 μm core/cladding).

Reviewer 2, Comment 4 — Number of participating spatial modes

The Yb fiber supports a few transverse modes, unlike the GRIN fiber that supports hundreds of modes. If the GRIN fiber was spliced on-axis with the output fibers from the coupler, then only the fundamental mode and a few HOM of the GRIN fiber will be excited. Can the Authors estimate the number of spatial modes involved in the cavity dynamics?

This is an important observation, and we agree that the number of *participating* modes is not the full GRIN support. The graded-index section supports of order hundreds of guided modes ($M \approx V^2/4$), whereas the Yb-doped section supports only a few transverse modes. On-axis coupling from the few-mode gain fiber preferentially excites the low-order GRIN subgroup, and the SLM-imposed phase structure diversifies this launch. The number of modes participating in the cavity dynamics is therefore set jointly by the few-mode Yb support and the SLM launch, and is of order *ten to a few tens* rather than the full hundreds-of-modes GRIN capacity. We have added this estimate, with the V -number scaling, to Supplementary Note S9.

Reviewer 2, Comment 5 — Single- versus multi-mode CW operation

In CW operation, the cavity operates on a single spatial mode, or it lases simultaneously on many spatial modes at different frequencies?

The cavity lases simultaneously on many spatial modes. Mode-differential gain saturation does not collapse the field to a single transverse mode; it redistributes power among the competing modes, and different transverse-mode groups generally lase at different longitudinal frequencies, so the steady state is multimode and multi-frequency. This multimode, multi-frequency operation is what supplies the high-dimensional, input-dependent feature space exploited for classification. We have clarified this in the Discussion, and the multi-frequency, multimode character is evidenced directly by the pump-power-dependent output spectra added under Reviewer 3, Comment 1 (Supplementary Note S10.4).

Added to the Discussion:

The cavity does not operate on a single transverse mode: gain saturation redistributes power among many competing spatial modes rather than selecting one, and distinct transverse-mode groups lase at distinct longitudinal frequencies, so steady-state operation is multimode and multi-frequency. This multimode lasing is the origin of the high-dimensional optical feature space.

Reviewer 3

The manuscript presents a continuous-wave Yb-doped multimode fiber laser cavity as a physical non-linear feature extractor for image classification. A spatial light modulator serves as a programmable cavity end mirror, the cavity is closed by a Sagnac loop, and the steady-state output intensity is read out with a linear Ridge classifier. The idea is interesting and the experimental effort is clearly substantial, but several issues require major revision.

We thank the Reviewer for the detailed and constructive assessment. We respond to each of the nine points below; several overlap with Reviewers 1 and 2 and are cross-referenced.

Reviewer 3, Comment 1 — Mechanistic controls (pump off, passive fiber, Kerr, spectra)

The main mechanistic claim is not yet convincingly established. The manuscript attributes the nonlinear mapping to gain saturation and multimode interference while neglecting Kerr effects, but the key controls are missing. The authors should show results with the pump turned off and, ideally, with the Yb-doped section replaced by a passive fiber. A simple estimate of the Kerr phase shift, together with output spectra as a function of pump power, would also help support the physical interpretation.

We agree these controls are essential and have addressed them in new Supplementary Note S10, with one important clarification regarding the pump-off control. The system is a laser whose only optical source is the pumped Yb-doped gain fiber: 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. Switching the pump off, or replacing the gain fiber with a passive fiber, therefore removes the only source of light in the system—the output is dark rather than linearly transformed, and there is no externally injected seed that could propagate through a passive cavity. A pump-off experimental control is thus not physically meaningful for this architecture. We instead establish the role of gain with (i) an analytical Kerr-phase estimate, (ii) a numerical gain-off control in the digital twin (where a seed field can be propagated with the gain term disabled).

Added to Methods:

The single-pass Kerr phase shift is $\Delta\phi_{\text{Kerr}} = (2\pi/\lambda)n_2IL \sim 10^{-5}$ rad at the intracavity intensities used here and is therefore neglected in Eq. (1); the nonlinear response is dominated by gain saturation. Because the pump is free-space coupled through a dichroic mirror and the resonator is formed between the SLM and the free-space path, switching off the pump (or replacing the gain fiber with a passive fiber) extinguishes the intracavity field entirely; the role of gain is therefore established by a numerical control in which the gain term is disabled while a seed field is propagated through the cavity (Supplementary Note S10).

New Supplementary Note S10 (full text):

Supplementary Note S10: 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 $\Delta\phi_{\text{Kerr}} = (2\pi/\lambda)n_2IL$, 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 $\Delta\phi_{\text{Kerr}} \sim 4 \times 10^{-5}$ 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 Eq. (1).

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.

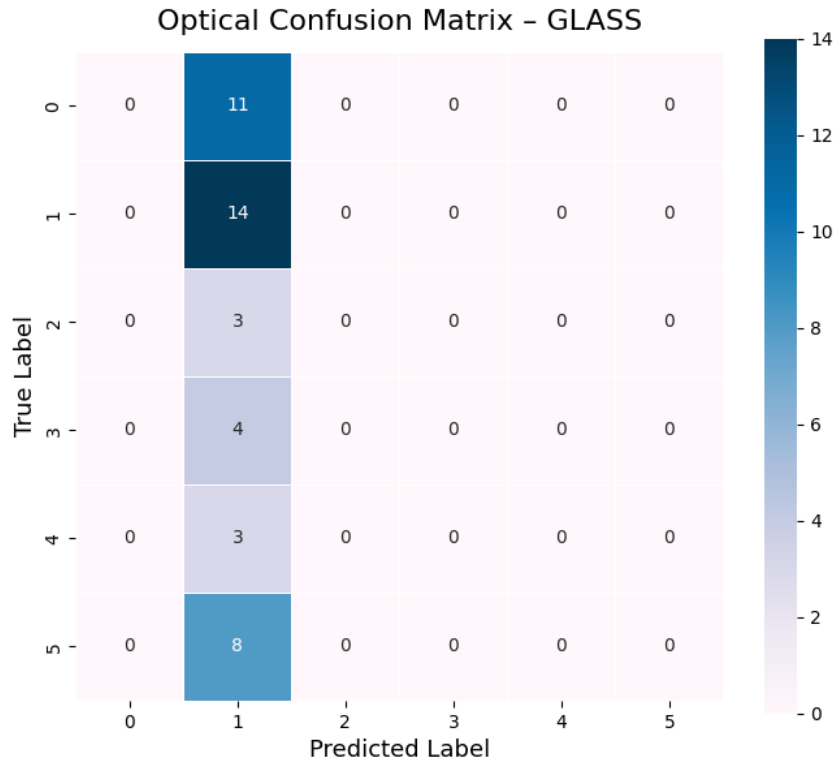


Figure 1: Confusion matrix obtained when the saturable-gain term is disabled ($g = 0$), leaving only linear multimode propagation and diffraction. The resulting classification accuracy decreases to approximately 32.56%, indicating that linear multimode mixing alone is insufficient to generate the nonlinear feature separation observed in the active cavity.

Reviewer 3, Comment 2 — Stronger baselines

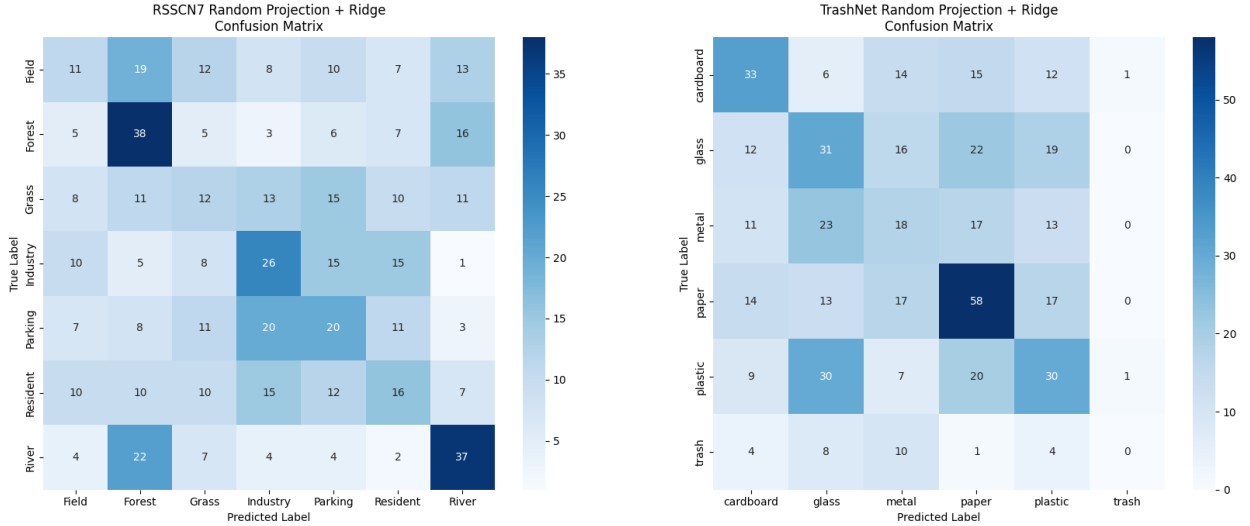
The baseline comparison is too weak. A linear classifier on raw pixels is not an adequate reference for a random-feature optical system. Stronger baselines are needed, in particular a matched random-projection-plus-Ridge baseline and a small digital model with comparable trainable complexity.

We agree and have added a matched Gaussian random-projection + Ridge baseline, applied to the raw datasets with the same Ridge readout and split. On RSSCN7 it reaches 28.57% accuracy

(28.57% balanced, 27.62% macro-F1) and on TrashNet 33.60% accuracy (28.89% balanced, 28.14% macro-F1)—far below the optically processed results (95.00% and 97.23%), confirming that a fixed linear random projection of comparable output dimensionality cannot account for the performance. The matched small-digital-model baseline is addressed under Reviewer 1, Comment 3 (no published 10^4 – 10^5 -parameter benchmark exists for these datasets; a LeNet-class model will be trained, see that action item). The new baseline confusion matrices are added to the Supplementary Material.

New Supplementary Figures 5 and 6:

Confusion matrices for the Gaussian random-projection + Ridge baselines, evaluated under the same pipeline as the optical features.



Supplementary Figure 5. Confusion matrix for the RSSCN7 Gaussian random-projection + Ridge baseline (28.57% accuracy).

Supplementary Figure 6. Confusion matrix for the TrashNet Gaussian random-projection + Ridge baseline (33.60% accuracy).

Reviewer 3, Comment 3 — OCT-MNIST clarification

The OCT MNIST result needs clarification. The reported accuracy is unusually high, and the use of a 4,000-image subset is not clearly explained. The authors should state explicitly which split was used, whether the reported metric is accuracy or AUC, and ideally verify the result using the standard MedMNIST evaluation pipeline.

We thank the Reviewer and have clarified the protocol. The OCT-MNIST experiment uses a class-balanced subset of 4,000 samples to keep the optical-acquisition load tractable while preserving class balance; the split is a stratified 80/20 train/test split with a fixed seed; and the reported metric is classification *accuracy* (not AUC). The test set therefore contains 800 images (200 per class), and the reported 98.87% corresponds to 791/800 correct (per-class metrics in Supplementary Table S3). This approach was adopted because the full MedMNIST dataset is substantially larger, making full optical acquisition and repeated experimental evaluation computationally and experimentally demanding.

Clarified in Results / Methods:

For OCT-MNIST, a class-balanced 4,000-image subset is used to limit the optical-acquisition load while preserving class balance. Results are reported as top-1 classification accuracy (not AUC) under a stratified 80/20 train/test split with a fixed seed; the test set comprises 800 images (200 per class).

Reviewer 3, Comment 4 — Statistical robustness

The statistical reporting is insufficient. All results appear to be single-point estimates. The authors should report performance over multiple random splits or seeds, and ideally across repeated experimental sessions, given the likely sensitivity to drift and calibration.

We agree and now report performance over multiple independent random splits (seeds = 10, 20, 30, 40). For optically processed TrashNet the system achieves $95.55 \pm 0.94\%$ accuracy, $95.44 \pm 1.11\%$ balanced accuracy, and $95.78 \pm 0.93\%$ macro-F1 across four independent splits. The small variance reflects the fact that the optical feature transformation is a fixed, deterministic physical map: the cavity is not retrained between splits, so split-to-split variation arises only from the data partition and the linear readout. This makes the single-session multi-split statistics representative, and no re-acquisition across sessions is required to establish stability for a fixed cavity.

Added to Results:

Because the optical transformation is a fixed, deterministic physical map (the cavity is not re-trained), we quantify variability over data partitions: across four independent random splits, optically processed TrashNet yields $95.55 \pm 0.94\%$ accuracy, $95.44 \pm 1.11\%$ balanced accuracy, and $95.78 \pm 0.93\%$ macro-F1, confirming that the result is not an artefact of a single partition.

Reviewer 3, Comment 5 — LDA fitted on training data only

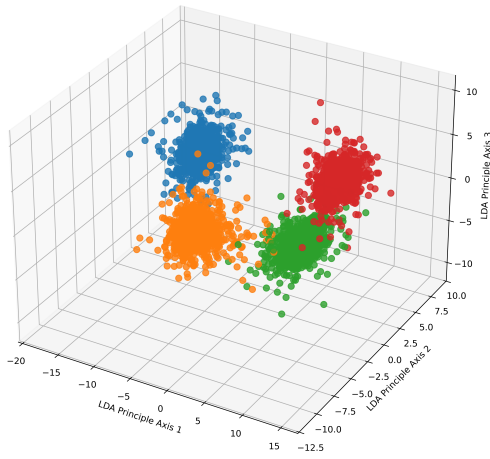
The LDA analysis in Fig. 5 may not be valid if the projection was fit using both training and test data. LDA should be fit on the training set only, with held-out test samples projected afterward.

We thank the Reviewer for catching this. We have recomputed all LDA visualisations by fitting the LDA transform on the training set only and projecting the held-out test samples afterward, using the same stratified split and fixed seed. The corrected projections retain the well-separated cluster structure, confirming that the separation is a property of the optical features rather than information leakage.

Updated Fig. 5 and its caption:

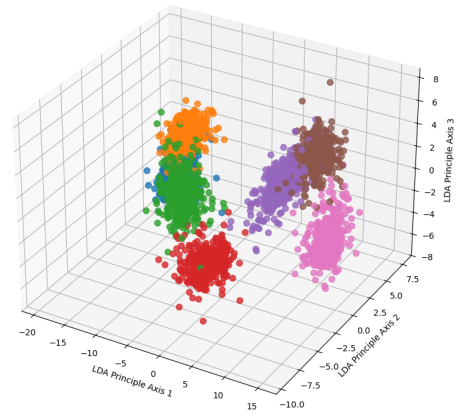
The LDA transform is fitted on the training set only; held-out test samples are projected afterward. The optically processed features form well-separated clusters under this train-only protocol, whereas raw inputs remain overlapping.

OCT-Mnist - Optically Processed - LDA



(a) Corrected OCT-MNIST LDA projection (LDA fitted on training data only; held-out test samples projected afterward).

RSSCN7 - Optically Processed - LDA

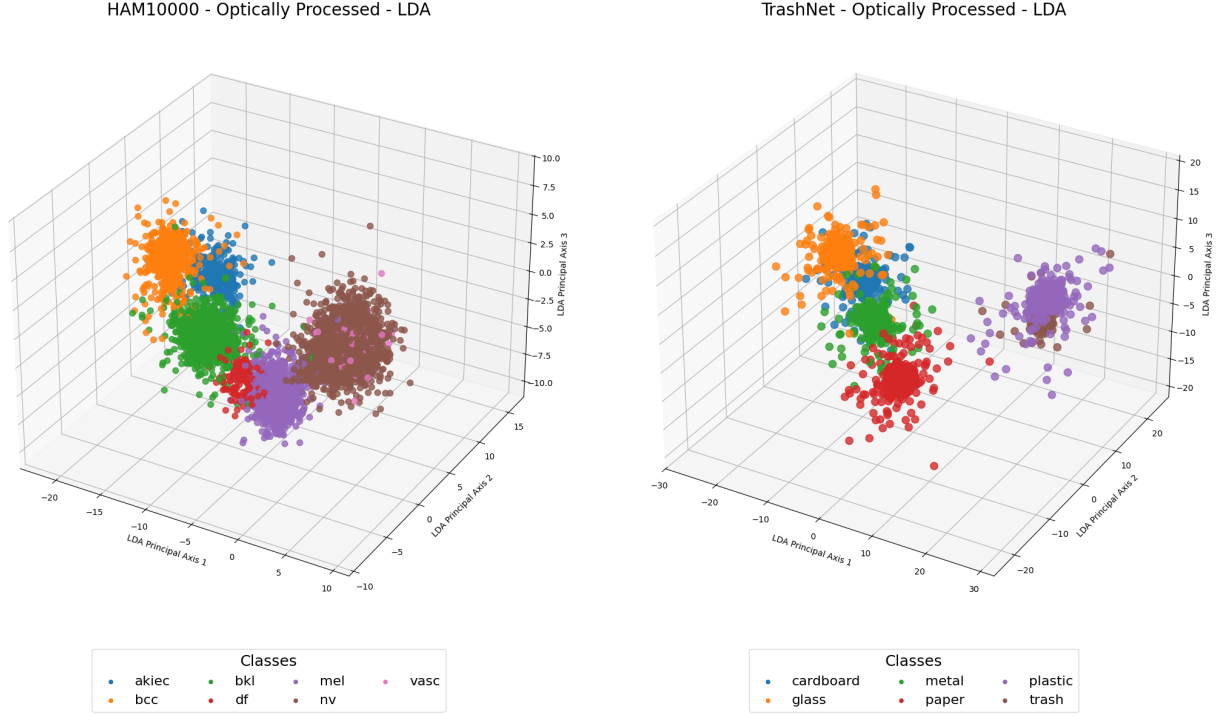


(b) Corrected RSSCN7 LDA projection (train-only fit, test samples projected afterward).

Figure 2: Corrected train-only LDA visualizations for the main-text datasets.

Updated Supplementary Figures 1 and 2 (corrected, train-only LDA):

The HAM10000 and TrashNet supplementary LDA panels (existing Supplementary Figs. 1–2) are regenerated under the same train-only fitting protocol.



(a) Corrected HAM10000 LDA projection (train-only fit, test samples projected afterward).

(b) Corrected TrashNet LDA projection (train-only fit, test samples projected afterward).

Figure 3: Corrected train-only LDA visualizations for the supplementary datasets.

Reviewer 3, Comment 6 — Class-balanced metrics for HAM10000

The HAM10000 result should be reported with class-balanced metrics. Since this dataset is imbalanced, overall accuracy alone is not sufficient. Balanced accuracy and per-class F1 scores should also be provided.

We agree and now report class-balanced metrics for HAM10000, computed directly from the confusion matrix of Fig. 4d. The overall accuracy of 85.27% is accompanied by a balanced accuracy (macro-recall) of 0.8849, a macro precision of 0.7006, and a macro-F1 of 0.7525; the lower macro precision and macro-F1 reflect reduced precision on the rare classes (*df*, *vasc*), exactly the imbalance effect the Reviewer notes. The full per-class precision, recall, and F1 are added as Supplementary Table S2, and the balanced OCT-MNIST metrics as Supplementary Table S3. We note openly that the reduced macro-F1 is driven primarily by weaker performance on the rarest classes (*vasc* precision 0.2385, *df* precision 0.5405): because the linear readout is trained without class weighting, minority classes exhibit reduced precision because the unweighted linear readout produces increased false-positive assignments for rare categories. Class-weighted or resampled readout training would likely improve minority-class performance at some cost to overall accuracy; we report the unweighted metrics for consistency with the other datasets and identify this rare-class behavior as a limitation of the current readout stage rather than of the optical feature extraction itself.

Added to Results:

For HAM10000, which is class-imbalanced, we report class-balanced metrics alongside overall accuracy: balanced accuracy (macro-recall) 0.8849, macro precision 0.7006, and macro-F1 0.7525 at an overall accuracy of 85.27%. Per-class precision, recall, and F1 are given in Supplementary Table S2. For OCT-MNIST the corresponding figures are balanced accuracy 0.9888, macro precision 0.9892, and macro-F1 0.9888 (Supplementary Table S3).

New Supplementary Table S2 (HAM10000, optically processed, per-class metrics from 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 S2: 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.

New Supplementary Table S3 (OCT-MNIST, optically processed, per-class metrics from 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 Table S3: Per-class precision, recall, and F1 for optically processed OCT-MNIST (800 test images, 200 per class), computed from the confusion matrix in Fig. 4b.

Reviewer 3, Comment 7 — Simulation–experiment gap

The gap between the simulation and experimental improvements is not discussed. In the UCI Glass simulation, the gain is modest, whereas in the experimental image datasets the reported gains over the raw baselines are much larger. The authors should comment on why the improvement differs so strongly across these cases.

We thank the Reviewer and have added this discussion. Three factors account for the difference, which we now state explicitly.

Added to the Discussion:

The improvement over the raw baseline is much larger for the experimental image datasets than for the UCI Glass simulation primarily because of the baseline floor. The high-dimensional image tasks have near-chance raw-pixel linear baselines ($\sim 18.0\text{--}24.0\%$), so any separable nonlinear feature map produces a large relative gain, whereas the low-dimensional Glass benchmark (nine features encoded as a simple stripe pattern) already has a comparatively strong raw-feature linear baseline, leaving

little headroom for enhancement. A secondary factor is that the experimental cavity samples additional physical degrees of freedom that are not represented in the reduced (2+1)D model such as full multimode and polarisation content, so the measured feature space is somewhat higher-dimensional than the simulated one. For these reasons the absolute gains should not be compared directly between the simulated and experimental settings.

Reviewer 3, Comment 8 — Literature on passive recurrent optical nonlinear processing

Important recent literature on passive recurrent optical nonlinear processing is missing and should be discussed, especially because it bears directly on the manuscript’s framing.

We thank the Reviewer for this important suggestion and have expanded the discussion to better position the proposed system within the broader context of recurrent optical and photonic reservoir computing architectures. In particular, we now discuss passive recurrent optical nonlinear processors including multimode waveguide ring-resonator reservoirs [13], recurrent speckle/scattering systems [12, 7], and structural processors that synthesise effective nonlinearity from linear optics [6]. We additionally include recent work on gain-controlled multimode-fibre photonic reservoir computing [20] and broader perspectives on machine learning in photonics [21].

The revised discussion emphasises that, unlike passive recurrent optical processors where nonlinearity is externally imposed or structurally synthesised, the present architecture combines recurrence with intrinsic active gain-mediated nonlinear cavity dynamics inside the laser cavity itself.

Added to the Introduction/Discussion:

Also, structural nonlinearities from linear optical elements arranged in a recurrent or multi-plane geometry, can be harnessed as a computing media [13, 12, 7, 6, 21]. In these optical processors the nonlinearity is either supplied by an external digital stage or emerges structurally from the recurrence of linear optical elements.

The use of these multimode complex field distributions were shown useful for optical computing in numerical simulations [20].

Reviewer 3, Comment 9 — Parameter-count comparison and better metrics

The parameter-count comparison in Table 1 is not fully fair. The manuscript compares readout parameters against the full parameter counts of CNN models. For this type of claim, energy per inference and throughput would be more informative metrics.

We agree. We retain the parameter-count column for context but reframe the comparison around energy per inference and throughput, reported in the new Supplementary Note S11 (power budget, energy per inference, TOPS and TOPS/W at the current operating point and projected under a faster modulator), with the protocol caveats stated. This is the same analysis added for Reviewer 1, Comment 4.

References

- [1] U. Teğın, M. Yıldırım, İ. Oğuz, C. Moser, and D. Psaltis, “Scalable optical learning operator,” *Nat. Comput. Sci.* **1**, 542–549 (2021).
- [2] B. U. Kesgin and U. Teğın, “Photonic neural networks at the edge of spatiotemporal chaos in multimode fibers,” *Nanophotonics* (2025).
- [3] B. Çarpınhoğlu and U. Teğın, “Genetically programmable optical random neural networks,” *Commun. Phys.* **8**, 349 (2025).

- [4] B. U. Kesgin, F. Yüce, and U. Teğın, “Fiber-based diffractive deep neural network,” *Opt. Lett.* **50**, 5254–5257 (2025).
- [5] X. Lin *et al.*, “All-optical machine learning using diffractive deep neural networks,” *Science* **361**, 1004–1008 (2018).
- [6] M. Yıldırım, N. U. Dinc, İ. Oğuz, D. Psaltis, and C. Moser, “Nonlinear processing with linear optics,” *Nat. Photonics* **18**, 1076–1082 (2024).
- [7] F. Xia *et al.*, “Nonlinear optical encoding enabled by recurrent linear scattering,” *Nat. Photonics* **18**, 1067–1075 (2024).
- [8] U. Paudel, M. Luengo-Kovac, J. Pilawa, T. J. Shaw, and G. C. Valley, “Classification of time-domain waveforms using a speckle-based optical reservoir computer,” *Opt. Express* **28**, 1225–1237 (2020).
- [9] L. G. Wright, D. N. Christodoulides, and F. W. Wise, “Controllable spatiotemporal nonlinear effects in multimode fibres,” *Nat. Photonics* **9**, 306–310 (2015).
- [10] L. G. Wright, W. H. Renninger, D. N. Christodoulides, and F. W. Wise, “Nonlinear multimode photonics: nonlinear optics with many degrees of freedom,” *Optica* **9**, 824–841 (2022).
- [11] S. Turitsyn *et al.*, “Modeling of CW Yb-doped fiber lasers with highly nonlinear cavity dynamics,” *Opt. Express* **19**, 8394–8405 (2011).
- [12] M. Rafayelyan, J. Dong, Y. Tan, F. Krzakala, and S. Gigan, “Large-scale optical reservoir computing for spatiotemporal chaotic systems prediction,” *Phys. Rev. X* **10**, 041037 (2020).
- [13] M. N. Ashner, U. Paudel, M. Luengo-Kovac, J. Pilawa, and G. C. Valley, “Photonic reservoir computer using speckle in multimode waveguide ring resonators,” *Opt. Express* **29**, 19262–19277 (2021).
- [14] H. Cao *et al.*, “Controlling light propagation in multimode fibers for imaging, spectroscopy, and beyond,” *Adv. Opt. Photonics* **15**, 524–612 (2023).
- [15] G. P. Agrawal, *Applications of Nonlinear Fiber Optics* (Academic Press, 2020).
- [16] 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.
- [17] P. Tschandl, C. Rosendahl, and H. Kittler, “The HAM10000 dataset: a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Sci. Data* **5**, 180161 (2018).
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE* **86**, 2278–2324 (1998).
- [19] B. Xu *et al.*, “Sparse transmission in diffractive deep neural networks via optical spiking neurons,” *Optica* **13**, 822–828 (2026).
- [20] G. Marcucci, L. Olivieri, and J. S. Toterogongora, “Optimising complexity and learning for photonic reservoir computing with gain-controlled multimode fibres,” *Front. Nanotechnol.* **7**, 1631564 (2025).
- [21] G. Genty, L. Salmela, J. M. Dudley *et al.*, “Machine learning and applications in ultrafast photonics,” *Nat. Photonics* **15**, 91–101 (2021).

Response to Reviewers

Multimode fiber laser cavities as nonlinear optical processors

We would like to thank all three Reviewers for carefully evaluating our work and for the constructive comments that have improved the quality and presentation of our study. We have responded to every comment and question in detail below. In this document the listed colours indicate the following:

- **Blue:** Comment of the Reviewer.
 - **Black:** Response of the Authors.
 - **Red:** Changes in the Manuscript.
 - **Purple:** Changes in the Supplementary Material.
-

Revision 2

We have performed the editorial revisions as requested in the attached document. We thank all Reviewers for their thorough evaluation of our study and for their constructive comments.

Reviewer 2

The Authors did not provide a point to point response letter to the previous reviewers' reports. Therefore, it is not possible to assess if the revision was made according to the previous reports.

We apologize for any confusion caused by the document hierarchy in the submission system. In the previous revision, our point-by-point response was uploaded under the heading "Author Cover Letter" rather than as a separately labelled response document. For clarity and ease of review, we reproduce the complete set of responses below.

Reviewer 3

The revised manuscript has addressed my previous concerns satisfactorily.

We thank the Reviewer for their positive evaluation and also for their constructive remarks at the previous round.

Revision 1

Reviewer 1

This manuscript presents an experimental demonstration of a multimode fiber laser cavity used as a nonlinear optical processor for image classification. The approach uses multimode interference and gain saturation to map input phase patterns into spatial intensity distributions, followed by a linear classifier. While the work is technically interesting and the experimental results are also

interesting, the overall contribution remains incremental and does not reach the level of novelty or physical advances expected for Communications Physics.

The main idea closely follows established areas of research in optical computing and multimode fiber processing. The use of nonlinear cavity dynamics as a feature mapping mechanism is well known, and the manuscript does not introduce a new concept or mechanism beyond this existing body of work. The work is an interesting implementation.

In addition, the manuscript provides limited physical description into the underlying processes. The role of multimode interference and gain saturation is described, but there is no deeper analysis of modal dynamics, scaling behavior or processing capacity. As a result, the work does not significantly advance our understanding of the physics governing these systems.

The reported performance is based on relatively simple datasets, and comparisons to neural networks are not carried out under equivalent conditions. This makes it difficult to understand the true significance of the results. Moreover, system advantages such as energy efficiency, latency, and scalability are not presented, which weakens the broader claims regarding practical optical computing.

Overall, the manuscript presents a very good experimental study, but the level of novelty and impact is more appropriate for a specialized optics journal rather than a high-impact physics journal.

We thank the Reviewer for the careful reading and for recognising the technical and experimental merit of the work. The Reviewer raises four connected concerns, which we address as Comments 1–4 below: (1) the conceptual novelty of the platform; (2) the absence of a deeper physical analysis of modal dynamics, scaling, and processing capacity; (3) the comparison to neural networks under equivalent conditions; and (4) the lack of system-level figures of merit. We have added a new theoretical treatment of gain-saturated modal competition (Supplementary Note S9) and a quantitative energy/throughput analysis (Supplementary Note S11) that directly target these points.

Reviewer 1, Comment 1 — Conceptual novelty

The use of nonlinear cavity dynamics as a feature mapping mechanism is well known, and the manuscript does not introduce a new concept or mechanism beyond this existing body of work.

We respectfully clarify the specific advance of this work, which we agree was under-stated in the original framing. Existing approaches to nonlinearity in multimode-optical computing fall into three families. (i) *Passive* speckle/scattering systems realise a fixed linear random projection of the launched field [8]. (ii) *Pulsed* nonlinear multimode systems obtain their nonlinearity from Kerr-mediated spatiotemporal dynamics at high peak intensity [9, 10]. (iii) *Structural* approaches synthesise an effective nonlinearity from linear elements arranged in a cavity or multi-plane geometry, for example multi-plane light conversion and recurrent linear scattering [6, 7]. Our platform is distinct from all three: it uses an *active*, continuous-wave laser cavity in which the nonlinearity is mode-competition-driven cross-gain saturation, and in which intracavity feedback makes the input-to-feature map *recurrent* rather than single-pass. The computational primitive is therefore a continuous-wave, low-power, gain-saturated nonlinear fixed-point map. We have sharpened the Introduction to state this distinction explicitly, supported by the new theory in Supplementary Note S9 and by the random-projection baselines (Reviewer 3, Comment 2), which confirm empirically that a fixed linear projection cannot reproduce the measured performance. The distinction also extends to our group’s own recent optical processors: unlike a passive, genetically-programmed optical random network [3] or pulsed multimode dynamics operated near the edge of spatiotemporal chaos [2], the present system is an *active, continuous-wave, gain-saturated* cavity operated in a stable, deterministic fixed-point regime. The concrete capability this enables, and which the passive and pulsed approaches do not, is a reproducible, low-power nonlinear feature map whose nonlinearity is supplied by the gain medium itself, so that a single fixed linear readout decodes stable class-separable features without any external nonlinearity, high-peak-power pulses, or trained

optical elements.

Added to the Introduction:

Also, structural nonlinearities from linear optical elements arranged in a recurrent or multi-plane geometry, can be harnessed as a computing media [13, 12, 7, 6, 21]. In these optical processors the nonlinearity is either supplied by an external digital stage or emerges structurally from the recurrence of linear optical elements.

Unlike previously reported, the optical feature extraction in this system occurs from gain-saturated fixed-point operation realised at low average power.

Reviewer 1, Comment 2 — Modal dynamics, scaling, and processing capacity

The role of multimode interference and gain saturation is described, but there is no deeper analysis of modal dynamics, scaling behavior or processing capacity.

We agree, and this comment prompted the most substantial addition to the revised manuscript: a dedicated theoretical treatment of the modal dynamics in Supplementary Note S9, and a corresponding paragraph in the main-text Discussion. The note (i) writes the cavity field in a modal basis, (ii) derives the cross-gain saturation that couples the modes and shows why it is a genuine, intrinsically multimode nonlinearity, (iii) frames the steady state as the fixed point of a recurrent nonlinear map, (iv) bounds the processing capacity and effective dimensionality, and (v) gives the scaling with mode number, including an estimate of the number of modes participating in the dynamics. The full text added to the manuscript and the Supplementary Material follows.

Added to the Methods:

The computational nonlinearity of the cavity originates in mode-differential gain saturation. The intracavity field is a superposition of guided modes,

$$A(x, y, z) = \sum_m a_m(z) \psi_m(x, y) e^{i\beta_m z} \quad (1)$$

whose phases advance at different rates so that multimode interference redistributes the transverse intensity in an input-dependent but linear, energy-conserving manner. The nonlinearity enters through the Yb-doped section, where all modes draw on a common, saturable inversion: the gain seen by mode m is expressed as

$$g_m = g_m^0 / (1 + P_{\text{sat}}^{-1} \sum_n \Gamma_{mn} P_n) \quad (2)$$

with modal power P_n and the overlap of modes m and n (Γ_{mn}) within the doped core. The off-diagonal overlaps couple the gain of every mode to the power of every other mode, so that a mode carrying more power suppresses the gain of its competitors. Cavity feedback iterates this nonlinear, mode-competitive map to an input-dependent steady state. The dimensionality of this representation is bounded by the modal content, up to $\sim M$ intensity channels and $\sim M(M+1)/2$ second-order interference features for M participating modes, and scales as $M \propto (a \text{NA}/\lambda)^2$. A complete derivation, together with an estimate of the number of participating modes, is given in Supplementary Note 9.

New Supplementary Note S9 (full text):

Supplementary Note S9: Modal gain competition, nonlinear 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 of order 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.

Reviewer 1, Comment 3 — Comparison to neural networks under equivalent conditions

The reported performance is based on relatively simple datasets, and comparisons to neural networks are not carried out under equivalent conditions.

We agree that the original comparison mixed evaluation protocols, and we have strengthened it in three ways. First, we add a matched Gaussian random-projection + Ridge baseline, evaluated with the same readout and split as the optical features (Reviewer 3, Comment 2): it reaches only 28.57% (RSSCN7) and 33.60% (TrashNet), confirming that a fixed linear projection of comparable output dimensionality cannot account for the performance. Second, we report the full set of published CNN baselines for RSSCN7 under the same 50%–50% split used for the multimode-laser (MML) system (new Supplementary Table S4), drawn from the benchmark study of Sun et al. [16]: AlexNet 91.85%, VGG16 93.57%, ResNet-50 93.64%, DCNN 77.00%, DAC 93.43%, TDFF 92.37%, TEX-Net-LF 94.00%, and GLNet variants 93.78–95.07%, all using 1.5×10^7 – 10^8 trainable parameters, against 94.21% for the MML system with 2,500–10,000 readout parameters. Third, we note that the platform belongs to the same family as other optical neural networks that offload feature extraction to optics with a compact electronic readout—diffractive deep neural networks [5], fiber-based diffractive networks [4], and multi-plane structural processors [6]—and that the MML reaches competitive accuracy with a trainable-parameter budget at the low end of this group.

We also note, for balance, that the comparison does not favour the MML system on every dataset. On the imbalanced, fine-grained HAM10000 task the deep CNNs reach higher accuracy (VGG16 96.05%, AlexNet 95.29%) than the MML system (85.22%, balanced accuracy 0.701), and we do not claim accuracy parity there. The defensible claim on that task is parameter and energy efficiency rather than absolute accuracy: the MML attains 85% with a 10^3 – 10^4 -parameter linear readout against 10^7 – 10^8 parameters for the CNNs. We have made this asymmetry explicit in the revised Discussion so that the cross-architecture comparison is stated even-handedly.

New Supplementary Table S4 (RSSCN7 model comparison, 50%–50% split):

Learner	Trainable parameters	RSSCN7 accuracy
MML computing (this work)	2,500–10,000	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 [Sun et al.]	deep ($\gtrsim 10^7$)	77.00%
TDFF [Sun et al.]	deep ($\gtrsim 10^7$)	92.37%
DAC [Sun et al.]	deep ($\gtrsim 10^7$)	93.43%
AlexNet [Sun et al.]	6.0×10^7	91.85%
VGG16 [Sun et al.]	1.38×10^8	93.57%
ResNet-50 [Sun et al.]	2.6×10^7	93.64%
TEX-Net-LF [Sun et al.]	deep ($\gtrsim 10^7$)	94.00%
GLNet (VGG16 backbone) [Sun et al.]	$> 1.38 \times 10^8$	95.07%

Supplementary Table S4: Comparison of trainable complexity and accuracy on RSSCN7 under a 50%–50% train–test split. CNN baselines are taken from Sun et al. [16]. The raw-pixel and random-projection rows are digital baselines of complexity comparable to the MML readout, evaluated under the identical pipeline.

Reviewer 1, Comment 4 — Energy efficiency, latency, and scalability

Moreover, system advantages such as energy efficiency, latency, and scalability are not presented, which weakens the broader claims regarding practical optical computing.

We have added a quantitative system-level analysis (new Supplementary Note S11) that reports a component-level power budget, the energy per inference and throughput at the current operating point, and the projected figures under a faster modulator, together with a comparison to a digital baseline. This also answers Reviewer 3, Comment 9. The analysis adopts the operations-counting framework of Xu et al. [19] used for diffractive optical processors so that the optical and digital figures are directly comparable. The full text added to the manuscript and the Supplementary Material follows.

Added to the Discussion:

At the current operating point the system consumes approximately 32.5 W in total (pump diode ≈ 5 W; SLM driver ≈ 24 W; host computer and camera ≈ 3.5 W). At the SLM-limited input frame rate of 60 Hz this corresponds to ≈ 0.54 J per inference and a throughput of ≈ 12.9 TOPS, or ≈ 0.40 TOPS/W (Supplementary Note S11). This figure is bottlenecked by the liquid-crystal SLM refresh rate and the digital host, not by the optical computation: the cavity performs the high-dimensional nonlinear feature transformation intrinsically during optical propagation and gain evolution, and the marginal optical cost per inference is negligible.

New Supplementary Note S11 (full text):

Supplementary Note S11: 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 S5. The dominant contribution is the liquid-crystal 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 $32.5 \text{ W}/60 \text{ Hz} \approx 0.54 \text{ J/inference}$.

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

Supplementary Table S5: Component-level electrical power budget for the multimode-fiber-laser proof-of-concept setup.

S11.2 Throughput and efficiency. At the current 60 Hz operating point, the throughput is $R_o \times 60 \approx 1.29 \times 10^{13} \text{ OPs/s} = 12.9 \text{ TOPS}$, i.e. $12.9/32.5 \approx 0.40 \text{ TOPS/W}$ (Supplementary Table S6). The gap to digital accelerators is dominated by the SLM-limited modulation rate, not by the optical computation. The cavity round-trip time is $t_{\text{rt}} = nL/c \approx 50 \text{ ns}$ for the $\sim 10 \text{ m}$ 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 $\approx 30 \text{ W}$, the throughput is $\approx 6.9 \times 10^3 \text{ TOPS}$ (6.9 POPS) at $\approx 230 \text{ TOPS/W}$, exceeding the NVIDIA A100 (1.56 TOPS/W) by more than two orders of magnitude.

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	$6.9 \times 10^3 \text{ TOPS}$	≈ 230
NVIDIA A100 (INT16)	—	—	312 TOPS	1.56

Supplementary Table S6: Throughput and energy efficiency of this work and a digital baseline. The

projected row assumes the modulator no longer limits the rate; the cavity buildup time sets the ultimate ceiling at $\sim 2 \times 10^5$ inferences per second.

Reviewer 2

There is significant interest in the study of optical analogue computers exploiting the parallel processing capability of optics, with the aim of reducing the energy requirements of conventional digital electronic processors, particularly for complex machine learning tasks. This manuscript presents an original and promising approach, based on a fiber laser cavity composed by multimode active and passive fibers. Spatial information is encoded by a means of a spatial light modulator (SLM) which provides a programmable mirror in the laser cavity. Different images encode phase patterns on the SLM produce different intensity distributions at the laser output, which are stable and well separable, demonstrating the capability of the device for image classification with relatively good accuracy. The main advantage of the implementation appears to be the strong reduction in the number of trainable parameters with respect to traditional neural network approaches. On the other hand, the processing speed is limited in this implementation by the low refresh rate of the liquid crystal based SLM. Overall, the demonstration is novel and convincing, and it represents a useful advance with respect to the state of the art to deserve publication. That said, I have some remarks on the numerical simulation of the laser cavity and on the operation of the fiber laser that would deserve additional clarifications by the Authors:

We thank the Reviewer for the positive overall assessment and for the detailed, technically precise questions, which we address point by point below.

Reviewer 2, Comment 1 — Index profiles and Sagnac-loop implementation

Eq.(1) is said to include diffraction and intensity dependent gain saturation, what about the refractive index profile of the two fibers that compose the cavity? Besides, it is not clear how the Sagnac loop was numerically implemented, did the Authors simulate a map of the main cavity and the loop until convergence was obtained? Simulation details and illustrative graphics should be added to the supplementary note S2.

We thank the Reviewer for this important point. The two fibers are modelled with their physical index profiles: a step-index profile for the Yb-doped gain fiber and a parabolic graded-index profile for the passive GRIN fiber. The transverse index distribution enters the diffraction operator $\hat{D}$ through the guiding term, so that the modal basis used in each section is that of the respective fiber rather than free space. The cavity is simulated as an iterated round-trip map, not a single pass: each round trip applies SLM phase encoding, propagation through the Yb-doped section with saturable gain, propagation through the GRIN section, and the 50/50 coupler/Sagnac loop, and the map is iterated until the relative change of the intracavity intensity profile between successive round trips falls below a fixed tolerance. We have added these details to Note S2.

Added to Supplementary Note S2:

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 SLM-encoded field, 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.

Reviewer 2, Comment 2 — Verification of the gain-saturation nonlinearity and laser operating parameters

It is claimed that nonlinearity is introduced by mode-differential gain saturation. Was that verified either experimentally and numerically? What is the saturation intensity? What is the intracavity intensity and the output power from the laser? In section 4.2 is written that experimental details are contained in note S6 but this note does not provide experimental details about the laser operation. What was the threshold power of the laser?

We thank the Reviewer for these precise questions. The saturation intensity of the Yb-doped fiber is $I_{\text{sat}} = h\nu/(\sigma\tau)$, with photon energy $h\nu$, effective emission cross-section σ , and upper-state lifetime τ ; using $\lambda = 1030$ nm, $\sigma \approx 2.5 \times 10^{-25}$ m² and $\tau \approx 0.84$ ms for Yb:silica [15], $I_{\text{sat}} \approx 9 \times 10^8$ W m⁻², corresponding to P_{sat} of order 0.1–0.3 W over a 20 μ m core. This places the operating point in the saturated-gain regime while remaining far below any Kerr-relevant intensity (Reviewer 3, Comment 1). The mode-differential nature of the saturation is verified numerically by the gain-off control of Supplementary Note S10: disabling the gain term collapses the feature separation toward the random-projection baseline. We have added the requested laser-operation parameters to Note S6.

Added to Supplementary Note S3:

The laser-operation parameters are as follows. The saturation intensity of the Yb-doped fiber is $I_{\text{sat}} = h\nu/(\sigma\tau) \approx 9 \times 10^8$ W m⁻² ($\lambda = 1030$ nm, $\sigma \approx 2.5 \times 10^{-25}$ m², $\tau \approx 0.84$ ms), equivalent to a saturation power of order 0.1–0.3 W over the 20 μ m gain-fiber core. These intensities are well below the Kerr-relevant scale (Supplementary Note S10), consistent with attributing the nonlinear feature mapping to gain saturation.

Reviewer 2, Comment 3 — Coupler and fiber specifications

What kind of 50/50 coupler was used between the active fiber and the GRIN fiber? Is it a multimode fiber coupler? What are the input and output fibers, graded index or step index, and what are their core/cladding sizes?

We thank the Reviewer and have added these specifications to the Methods. The 50/50 coupler is a multimode fiber coupler with graded-index multimode pigtails matched to the cavity GRIN fiber. The gain section is a step-index Yb-doped fiber (20/125 μ m core/cladding); the passive section is a graded-index multimode fiber (50/125 μ m core/cladding).

Added to Methods (Experimental setup):

The cavity uses a 50/50 multimode fiber coupler with graded-index multimode pigtails (50/125 μ m) matched to the passive cavity fiber. The gain section is a step-index Yb-doped fiber (20/125 μ m core/cladding); the passive section is a graded-index multimode fiber (50/125 μ m core/cladding).

Reviewer 2, Comment 4 — Number of participating spatial modes

The Yb fiber supports a few transverse modes, unlike the GRIN fiber that supports hundreds of modes. If the GRIN fiber was spliced on-axis with the output fibers from the coupler, then only the fundamental mode and a few HOM of the GRIN fiber will be excited. Can the Authors estimate the number of spatial modes involved in the cavity dynamics?

This is an important observation, and we agree that the number of *participating* modes is not the full GRIN support. The graded-index section supports of order hundreds of guided modes ($M \approx V^2/4$), whereas the Yb-doped section supports only a few transverse modes. On-axis coupling from the few-mode gain fiber preferentially excites the low-order GRIN subgroup, and the SLM-imposed phase structure diversifies this launch. The number of modes participating in the cavity dynamics is therefore set jointly by the few-mode Yb support and the SLM launch, and is of order *ten to a few tens* rather than the full hundreds-of-modes GRIN capacity. We have added this estimate, with the V -number scaling, to Supplementary Note S9.

Reviewer 2, Comment 5 — Single- versus multi-mode CW operation

In CW operation, the cavity operates on a single spatial mode, or it lases simultaneously on many spatial modes at different frequencies?

The cavity lases simultaneously on many spatial modes. Mode-differential gain saturation does not collapse the field to a single transverse mode; it redistributes power among the competing modes, and different transverse-mode groups generally lase at different longitudinal frequencies, so the steady state is multimode and multi-frequency. This multimode, multi-frequency operation is what supplies the high-dimensional, input-dependent feature space exploited for classification. We have clarified this in the Discussion, and the multi-frequency, multimode character is evidenced directly by the pump-power-dependent output spectra added under Reviewer 3, Comment 1 (Supplementary Note S10.4).

Added to the Discussion:

The cavity does not operate on a single transverse mode: gain saturation redistributes power among many competing spatial modes rather than selecting one, and distinct transverse-mode groups lase at distinct longitudinal frequencies, so steady-state operation is multimode and multi-frequency. This multimode lasing is the origin of the high-dimensional optical feature space.

Reviewer 3

The manuscript presents a continuous-wave Yb-doped multimode fiber laser cavity as a physical non-linear feature extractor for image classification. A spatial light modulator serves as a programmable cavity end mirror, the cavity is closed by a Sagnac loop, and the steady-state output intensity is read out with a linear Ridge classifier. The idea is interesting and the experimental effort is clearly substantial, but several issues require major revision.

We thank the Reviewer for the detailed and constructive assessment. We respond to each of the nine points below; several overlap with Reviewers 1 and 2 and are cross-referenced.

Reviewer 3, Comment 1 — Mechanistic controls (pump off, passive fiber, Kerr, spectra)

The main mechanistic claim is not yet convincingly established. The manuscript attributes the nonlinear mapping to gain saturation and multimode interference while neglecting Kerr effects, but the key controls are missing. The authors should show results with the pump turned off and, ideally, with the Yb-doped section replaced by a passive fiber. A simple estimate of the Kerr phase shift, together with output spectra as a function of pump power, would also help support the physical interpretation.

We agree these controls are essential and have addressed them in new Supplementary Note S10, with one important clarification regarding the pump-off control. The system is a laser whose only

optical source is the pumped Yb-doped gain fiber: 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. Switching the pump off, or replacing the gain fiber with a passive fiber, therefore removes the only source of light in the system—the output is dark rather than linearly transformed, and there is no externally injected seed that could propagate through a passive cavity. A pump-off experimental control is thus not physically meaningful for this architecture. We instead establish the role of gain with (i) an analytical Kerr-phase estimate, (ii) a numerical gain-off control in the digital twin (where a seed field can be propagated with the gain term disabled).

Added to Methods:

The single-pass Kerr phase shift is $\Delta\phi_{\text{Kerr}} = (2\pi/\lambda)n_2IL \sim 10^{-5}$ rad at the intracavity intensities used here and is therefore neglected in Eq. (1); the nonlinear response is dominated by gain saturation. Because the pump is free-space coupled through a dichroic mirror and the resonator is formed between the SLM and the free-space path, switching off the pump (or replacing the gain fiber with a passive fiber) extinguishes the intracavity field entirely; the role of gain is therefore established by a numerical control in which the gain term is disabled while a seed field is propagated through the cavity (Supplementary Note S10).

New Supplementary Note S10 (full text):

Supplementary Note S10: 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 $\Delta\phi_{\text{Kerr}} = (2\pi/\lambda)n_2IL$, 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 $\Delta\phi_{\text{Kerr}} \sim 4 \times 10^{-5}$ 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 Eq. (1).

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.

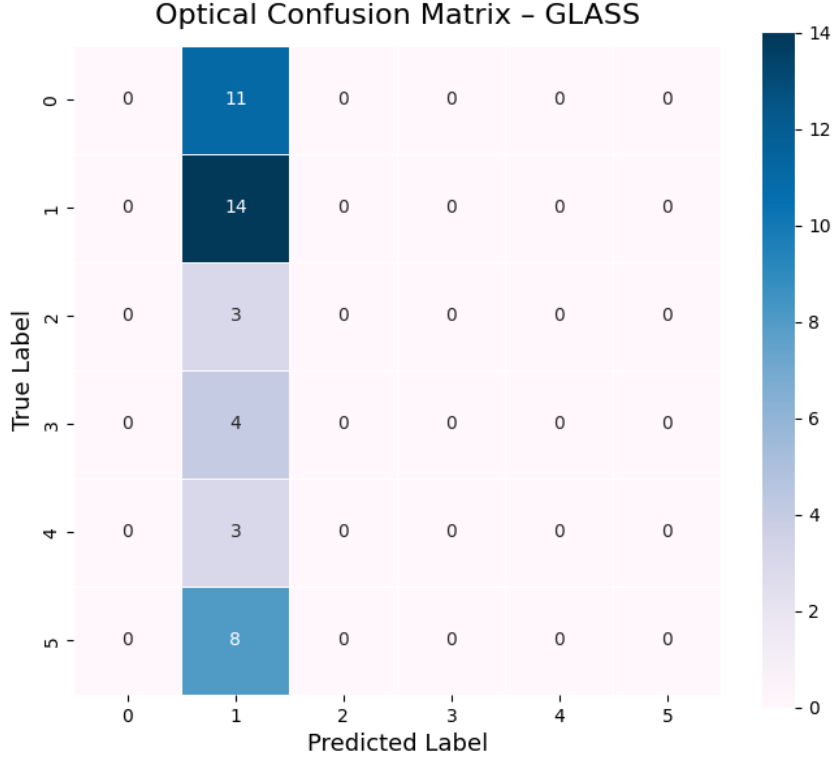


Figure 1: Confusion matrix obtained when the saturable-gain term is disabled ($g = 0$), leaving only linear multimode propagation and diffraction. The resulting classification accuracy decreases to approximately 32.56%, indicating that linear multimode mixing alone is insufficient to generate the nonlinear feature separation observed in the active cavity.

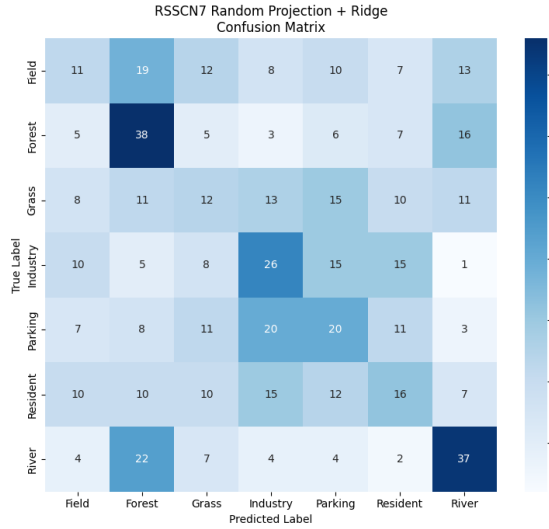
Reviewer 3, Comment 2 — Stronger baselines

The baseline comparison is too weak. A linear classifier on raw pixels is not an adequate reference for a random-feature optical system. Stronger baselines are needed, in particular a matched random-projection-plus-Ridge baseline and a small digital model with comparable trainable complexity.

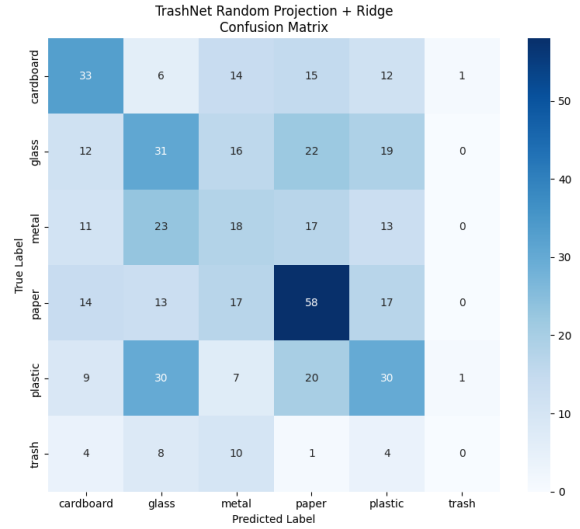
We agree and have added a matched Gaussian random-projection + Ridge baseline, applied to the raw datasets with the same Ridge readout and split. On RSSCN7 it reaches 28.57% accuracy (28.57% balanced, 27.62% macro-F1) and on TrashNet 33.60% accuracy (28.89% balanced, 28.14% macro-F1)—far below the optically processed results (95.00% and 97.23%), confirming that a fixed linear random projection of comparable output dimensionality cannot account for the performance. The matched small-digital-model baseline is addressed under Reviewer 1, Comment 3 (no published 10^4 – 10^5 -parameter benchmark exists for these datasets; a LeNet-class model will be trained, see that action item). The new baseline confusion matrices are added to the Supplementary Material.

New Supplementary Figures 5 and 6:

Confusion matrices for the Gaussian random-projection + Ridge baselines, evaluated under the same pipeline as the optical features.



Supplementary Figure 5. Confusion matrix for the RSSCN7 Gaussian random-projection + Ridge baseline (28.57% accuracy).



Supplementary Figure 6. Confusion matrix for the TrashNet Gaussian random-projection + Ridge baseline (33.60% accuracy).

Reviewer 3, Comment 3 — OCT-MNIST clarification

The OCT MNIST result needs clarification. The reported accuracy is unusually high, and the use of a 4,000-image subset is not clearly explained. The authors should state explicitly which split was used, whether the reported metric is accuracy or AUC, and ideally verify the result using the standard MedMNIST evaluation pipeline.

We thank the Reviewer and have clarified the protocol. The OCT-MNIST experiment uses a class-balanced subset of 4,000 samples to keep the optical-acquisition load tractable while preserving class balance; the split is a stratified 80/20 train/test split with a fixed seed; and the reported metric is classification *accuracy* (not AUC). The test set therefore contains 800 images (200 per class), and the reported 98.87% corresponds to 791/800 correct (per-class metrics in Supplementary Table S3). This approach was adopted because the full MedMNIST dataset is substantially larger, making full optical acquisition and repeated experimental evaluation computationally and experimentally demanding.

Clarified in Results / Methods:

For OCT-MNIST, a class-balanced 4,000-image subset is used to limit the optical-acquisition load while preserving class balance. Results are reported as top-1 classification accuracy (not AUC) under a stratified 80/20 train/test split with a fixed seed; the test set comprises 800 images (200 per class).

Reviewer 3, Comment 4 — Statistical robustness

The statistical reporting is insufficient. All results appear to be single-point estimates. The authors should report performance over multiple random splits or seeds, and ideally across repeated experimental sessions, given the likely sensitivity to drift and calibration.

We agree and now report performance over multiple independent random splits (seeds = 10, 20, 30, 40). For optically processed TrashNet the system achieves $95.55 \pm 0.94\%$ accuracy, $95.44 \pm 1.11\%$

balanced accuracy, and $95.78 \pm 0.93\%$ macro-F1 across four independent splits. The small variance reflects the fact that the optical feature transformation is a fixed, deterministic physical map: the cavity is not retrained between splits, so split-to-split variation arises only from the data partition and the linear readout. This makes the single-session multi-split statistics representative, and no re-acquisition across sessions is required to establish stability for a fixed cavity.

Added to Results:

Because the optical transformation is a fixed, deterministic physical map (the cavity is not re-trained), we quantify variability over data partitions: across four independent random splits, optically processed TrashNet yields $95.55 \pm 0.94\%$ accuracy, $95.44 \pm 1.11\%$ balanced accuracy, and $95.78 \pm 0.93\%$ macro-F1, confirming that the result is not an artefact of a single partition.

Reviewer 3, Comment 5 — LDA fitted on training data only

The LDA analysis in Fig. 5 may not be valid if the projection was fit using both training and test data. LDA should be fit on the training set only, with held-out test samples projected afterward.

We thank the Reviewer for catching this. We have recomputed all LDA visualisations by fitting the LDA transform on the training set only and projecting the held-out test samples afterward, using the same stratified split and fixed seed. The corrected projections retain the well-separated cluster structure, confirming that the separation is a property of the optical features rather than information leakage.

Updated Fig. 5 and its caption:

The LDA transform is fitted on the training set only; held-out test samples are projected afterward. The optically processed features form well-separated clusters under this train-only protocol, whereas raw inputs remain overlapping.

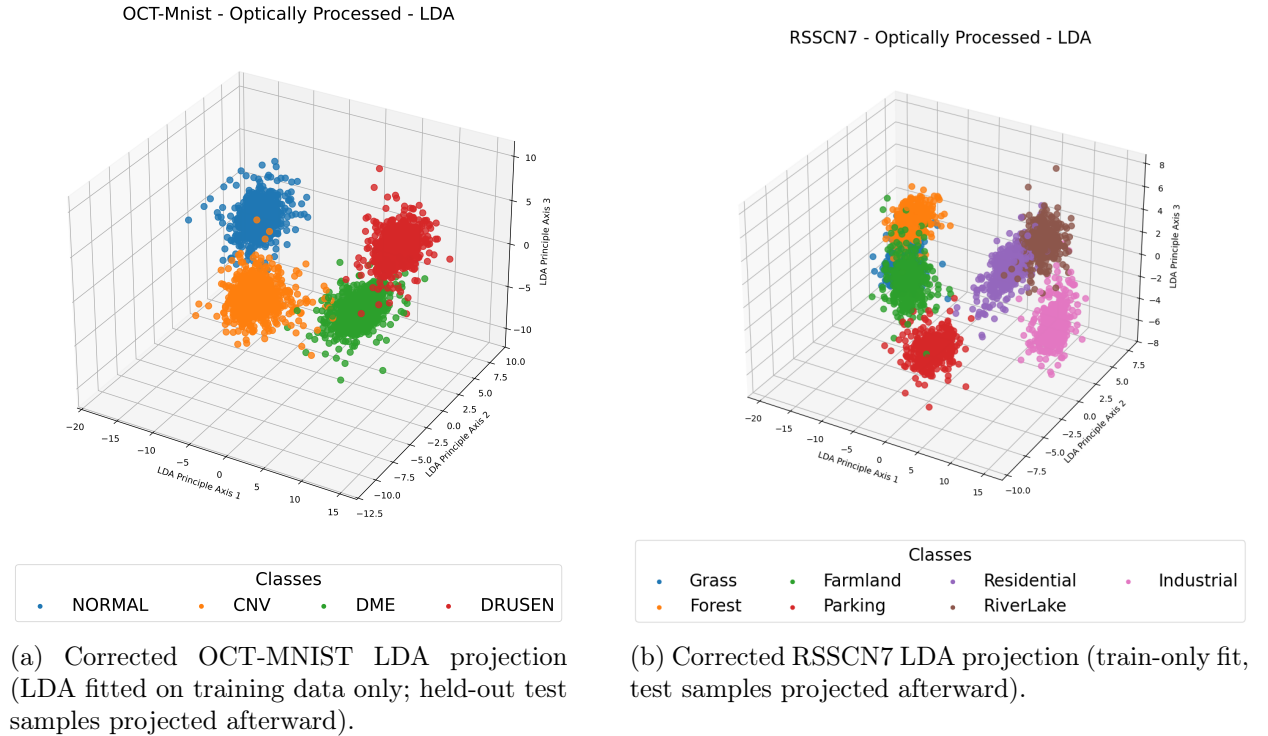
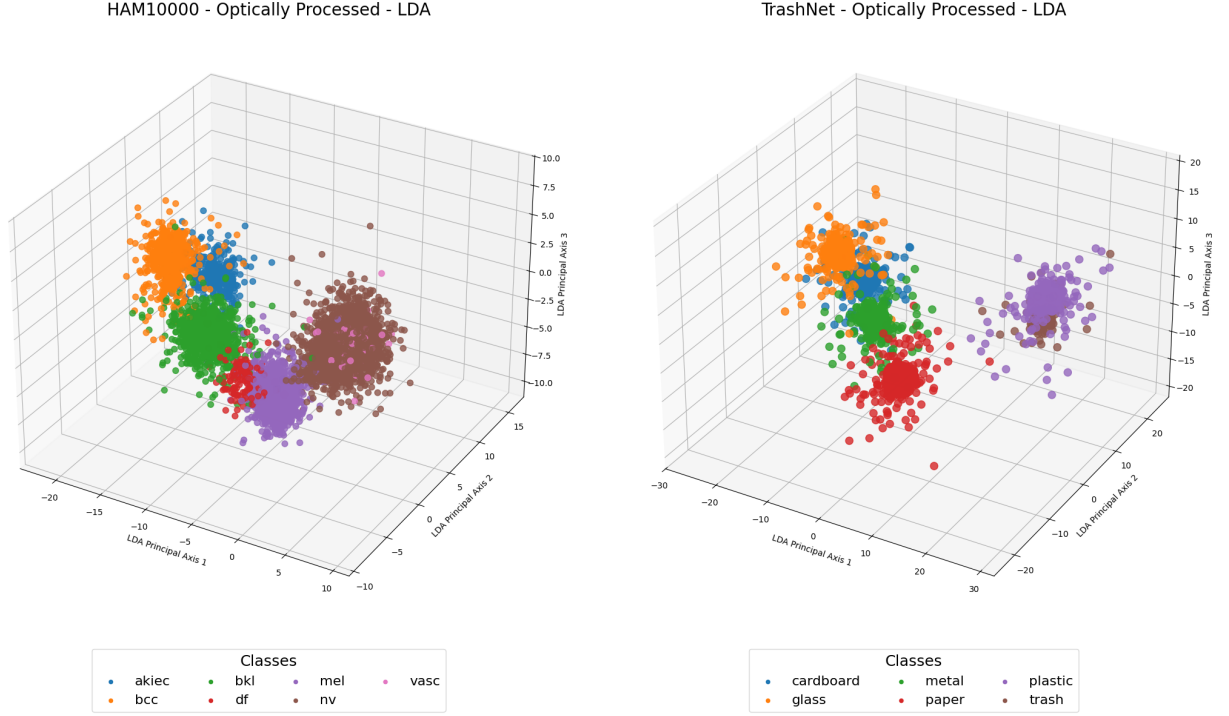


Figure 2: Corrected train-only LDA visualizations for the main-text datasets.

Updated Supplementary Figures 1 and 2 (corrected, train-only LDA):

The HAM10000 and TrashNet supplementary LDA panels (existing Supplementary Figs. 1–2) are regenerated under the same train-only fitting protocol.



(a) Corrected HAM10000 LDA projection (train-only fit, test samples projected afterward).

(b) Corrected TrashNet LDA projection (train-only fit, test samples projected afterward).

Figure 3: Corrected train-only LDA visualizations for the supplementary datasets.

Reviewer 3, Comment 6 — Class-balanced metrics for HAM10000

The HAM10000 result should be reported with class-balanced metrics. Since this dataset is imbalanced, overall accuracy alone is not sufficient. Balanced accuracy and per-class F1 scores should also be provided.

We agree and now report class-balanced metrics for HAM10000, computed directly from the confusion matrix of Fig. 4d. The overall accuracy of 85.27% is accompanied by a balanced accuracy (macro-recall) of 0.8849, a macro precision of 0.7006, and a macro-F1 of 0.7525; the lower macro precision and macro-F1 reflect reduced precision on the rare classes (*df*, *vasc*), exactly the imbalance effect the Reviewer notes. The full per-class precision, recall, and F1 are added as Supplementary Table S2, and the balanced OCT-MNIST metrics as Supplementary Table S3. We note openly that the reduced macro-F1 is driven primarily by weaker performance on the rarest classes (*vasc* precision 0.2385, *df* precision 0.5405): because the linear readout is trained without class weighting, minority classes exhibit reduced precision because the unweighted linear readout produces increased false-positive assignments for rare categories. Class-weighted or resampled readout training would likely improve minority-class performance at some cost to overall accuracy; we report the unweighted metrics for consistency with the other datasets and identify this rare-class behavior as a limitation of the current readout stage rather than of the optical feature extraction itself.

Added to Results:

For HAM10000, which is class-imbalanced, we report class-balanced metrics alongside overall accuracy: balanced accuracy (macro-recall) 0.8849, macro precision 0.7006, and macro-F1 0.7525 at an overall accuracy of 85.27%. Per-class precision, recall, and F1 are given in Supplementary Table S2. For OCT-MNIST the corresponding figures are balanced accuracy 0.9888, macro precision 0.9892, and macro-F1 0.9888 (Supplementary Table S3).

New Supplementary Table S2 (HAM10000, optically processed, per-class metrics from 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 S2: 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.

New Supplementary Table S3 (OCT-MNIST, optically processed, per-class metrics from 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 Table S3: Per-class precision, recall, and F1 for optically processed OCT-MNIST (800 test images, 200 per class), computed from the confusion matrix in Fig. 4b.

Reviewer 3, Comment 7 — Simulation–experiment gap

The gap between the simulation and experimental improvements is not discussed. In the UCI Glass simulation, the gain is modest, whereas in the experimental image datasets the reported gains over the raw baselines are much larger. The authors should comment on why the improvement differs so strongly across these cases.

We thank the Reviewer and have added this discussion. Three factors account for the difference, which we now state explicitly.

Added to the Discussion:

The improvement over the raw baseline is much larger for the experimental image datasets than for the UCI Glass simulation primarily because of the baseline floor. The high-dimensional image tasks have near-chance raw-pixel linear baselines ($\sim 18.0\text{--}24.0\%$), so any separable nonlinear feature map produces a large relative gain, whereas the low-dimensional Glass benchmark (nine features encoded as a simple stripe pattern) already has a comparatively strong raw-feature linear baseline, leaving

little headroom for enhancement. A secondary factor is that the experimental cavity samples additional physical degrees of freedom that are not represented in the reduced (2+1)D model such as full multimode and polarisation content, so the measured feature space is somewhat higher-dimensional than the simulated one. For these reasons the absolute gains should not be compared directly between the simulated and experimental settings.

Reviewer 3, Comment 8 — Literature on passive recurrent optical nonlinear processing

Important recent literature on passive recurrent optical nonlinear processing is missing and should be discussed, especially because it bears directly on the manuscript’s framing.

We thank the Reviewer for this important suggestion and have expanded the discussion to better position the proposed system within the broader context of recurrent optical and photonic reservoir computing architectures. In particular, we now discuss passive recurrent optical nonlinear processors including multimode waveguide ring-resonator reservoirs [13], recurrent speckle/scattering systems [12, 7], and structural processors that synthesise effective nonlinearity from linear optics [6]. We additionally include recent work on gain-controlled multimode-fibre photonic reservoir computing [20] and broader perspectives on machine learning in photonics [21].

The revised discussion emphasises that, unlike passive recurrent optical processors where nonlinearity is externally imposed or structurally synthesised, the present architecture combines recurrence with intrinsic active gain-mediated nonlinear cavity dynamics inside the laser cavity itself.

Added to the Introduction/Discussion:

Also, structural nonlinearities from linear optical elements arranged in a recurrent or multi-plane geometry, can be harnessed as a computing media [13, 12, 7, 6, 21]. In these optical processors the nonlinearity is either supplied by an external digital stage or emerges structurally from the recurrence of linear optical elements.

The use of these multimode complex field distributions were shown useful for optical computing in numerical simulations [20].

Reviewer 3, Comment 9 — Parameter-count comparison and better metrics

The parameter-count comparison in Table 1 is not fully fair. The manuscript compares readout parameters against the full parameter counts of CNN models. For this type of claim, energy per inference and throughput would be more informative metrics.

We agree. We retain the parameter-count column for context but reframe the comparison around energy per inference and throughput, reported in the new Supplementary Note S11 (power budget, energy per inference, TOPS and TOPS/W at the current operating point and projected under a faster modulator), with the protocol caveats stated. This is the same analysis added for Reviewer 1, Comment 4.

References

- [1] U. Teğın, M. Yıldırım, İ. Oğuz, C. Moser, and D. Psaltis, “Scalable optical learning operator,” *Nat. Comput. Sci.* **1**, 542–549 (2021).
- [2] B. U. Kesgin and U. Teğın, “Photonic neural networks at the edge of spatiotemporal chaos in multimode fibers,” *Nanophotonics* (2025).
- [3] B. Çarpınhoğlu and U. Teğın, “Genetically programmable optical random neural networks,” *Commun. Phys.* **8**, 349 (2025).

- [4] B. U. Kesgin, F. Yüce, and U. Teğın, “Fiber-based diffractive deep neural network,” *Opt. Lett.* **50**, 5254–5257 (2025).
- [5] X. Lin *et al.*, “All-optical machine learning using diffractive deep neural networks,” *Science* **361**, 1004–1008 (2018).
- [6] M. Yıldırım, N. U. Dinc, İ. Oğuz, D. Psaltis, and C. Moser, “Nonlinear processing with linear optics,” *Nat. Photonics* **18**, 1076–1082 (2024).
- [7] F. Xia *et al.*, “Nonlinear optical encoding enabled by recurrent linear scattering,” *Nat. Photonics* **18**, 1067–1075 (2024).
- [8] U. Paudel, M. Luengo-Kovac, J. Pilawa, T. J. Shaw, and G. C. Valley, “Classification of time-domain waveforms using a speckle-based optical reservoir computer,” *Opt. Express* **28**, 1225–1237 (2020).
- [9] L. G. Wright, D. N. Christodoulides, and F. W. Wise, “Controllable spatiotemporal nonlinear effects in multimode fibres,” *Nat. Photonics* **9**, 306–310 (2015).
- [10] L. G. Wright, W. H. Renninger, D. N. Christodoulides, and F. W. Wise, “Nonlinear multimode photonics: nonlinear optics with many degrees of freedom,” *Optica* **9**, 824–841 (2022).
- [11] S. Turitsyn *et al.*, “Modeling of CW Yb-doped fiber lasers with highly nonlinear cavity dynamics,” *Opt. Express* **19**, 8394–8405 (2011).
- [12] M. Rafayelyan, J. Dong, Y. Tan, F. Krzakala, and S. Gigan, “Large-scale optical reservoir computing for spatiotemporal chaotic systems prediction,” *Phys. Rev. X* **10**, 041037 (2020).
- [13] M. N. Ashner, U. Paudel, M. Luengo-Kovac, J. Pilawa, and G. C. Valley, “Photonic reservoir computer using speckle in multimode waveguide ring resonators,” *Opt. Express* **29**, 19262–19277 (2021).
- [14] H. Cao *et al.*, “Controlling light propagation in multimode fibers for imaging, spectroscopy, and beyond,” *Adv. Opt. Photonics* **15**, 524–612 (2023).
- [15] G. P. Agrawal, *Applications of Nonlinear Fiber Optics* (Academic Press, 2020).
- [16] 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.
- [17] P. Tschandl, C. Rosendahl, and H. Kittler, “The HAM10000 dataset: a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Sci. Data* **5**, 180161 (2018).
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE* **86**, 2278–2324 (1998).
- [19] B. Xu *et al.*, “Sparse transmission in diffractive deep neural networks via optical spiking neurons,” *Optica* **13**, 822–828 (2026).
- [20] G. Marcucci, L. Olivieri, and J. S. Toterogongora, “Optimising complexity and learning for photonic reservoir computing with gain-controlled multimode fibres,” *Front. Nanotechnol.* **7**, 1631564 (2025).
- [21] G. Genty, L. Salmela, J. M. Dudley *et al.*, “Machine learning and applications in ultrafast photonics,” *Nat. Photonics* **15**, 91–101 (2021).