Scalable optical learning operator

Uğur Teğın^{1,2}✉, Mustafa Yıldırım², İlker Oğuz^{1,2}, Christophe Moser² and Demetri Psaltis¹

Today's heavy machine learning tasks are fueled by large datasets. Computing is performed with power-hungry processors whose performance is ultimately limited by the data transfer to and from memory. Optics is a powerful means of communicating and processing information, and there is currently intense interest in optical information processing for realizing high-speed computations. Here we present and experimentally demonstrate an optical computing framework called scalable optical learning operator, which is based on spatiotemporal effects in multimode fibers for a range of learning tasks including classifying COVID-19 X-ray lung images, speech recognition and predicting age from images of faces. The presented framework addresses the energy scaling problem of existing systems without compromising speed. We leverage simultaneous, linear and nonlinear interaction of spatial modes as a computation engine. We numerically and experimentally show the ability of the method to execute several different tasks with accuracy comparable with a digital implementation.

Early optical computers were used to calculate linear operations such as Fourier transforms and correlations. They found applications in pattern recognition and synthetic aperture radar^{1,2}; however, with the advent of modern very-large-scale integration technology and efficient algorithms, digital signal processing based on silicon circuits became so fast and parallel that the analog optical computation including the input and output electronic overhead became obsolete. Digital optical computing, which combined nonlinear optical switches³ with linear optical interconnections⁴ that replaced wires, was then intensely pursued in the 1980s. Optical interconnections can be advantageous in terms of power consumption⁵; however, in an all-optical implementation this advantage is counter-balanced by the power inefficiency and large size of optical switches compared with the electronic ones. All-optical digital computers are therefore not yet competitive. Optics has also been used for the implementation of nonlinear computations that are not based on Boolean logic, such as the optical implementation of neural networks^{6,7}. In principle, the dense connectivity of neural networks and their relative robustness against noise and device imperfections renders them as a promising area for optical computing.

Interest in optically implemented neural networks has intensified in recent years, partially due to the large size of databases that need to be managed, stressing the capabilities of existing digital, electronic computers. Several promising approaches are being investigated and they are summarized in a recent review article⁸. The key challenge in designing a viable optical computer (including a neural one) is to combine the linear part of the system from where the competitive edge of optics derives, with nonlinear elements and input–output interfaces, while maintaining the speed and power efficiency of the optical interconnections.

The solution we propose and demonstrate in this paper is the combination of the linear and nonlinear parts of the optical system in a shared volume confined in a multimode fiber (MMF). The principal advantage of this approach is the combination of the three-dimensional connectivity of optics with the long interaction length and lateral confinement afforded by the fiber, which makes it possible to realize optical nonlinearities at relatively low optical power. At the same time, the large number of spatial modes that can be densely supported in a MMF maintains the traditional high parallelism feature of optics while maintaining

a compact form factor. Finally, with the availability of megapixel spatial light modulators (SLMs) and cameras, the two-dimensional input and output interfaces to the MMF can sustain a large information processing throughput. We refer to the proposed method as SOLO (scalable optical learning operator) for the remainder of this paper.

A schematic diagram of the MMF processing element is shown in Fig. 1. The data to be processed are entered through the two-dimensional SLM on the left. At sufficiently high-illumination peak power, the light from a pulsed light source is nonlinearly transformed as it propagates through the fiber, and the result of the computation is projected on the two-dimensional camera. Given the properties of the fiber and the laser source, the input–output operation performed by the MMF is fixed and highly nonlinear. We implement a reconfigurable processor by combining the fixed nonlinear MMF mapping in the optical domain with a single-layer digital neural network (decision layer) trained to recognize the output recorded on the camera using a large dataset of input–output pairs. For instance, we used this system to diagnose COVID-19 from X-ray images of lungs at high accuracy (83%). A large database of X-ray images of lungs with various diseases including COVID-19 was used to train the single-layer network that classifies the representation of the lungs produced at the output camera. The notion of combining a complex, fixed mapping with a simpler programmable processor to realize a powerful overall system—including the optical implementation of such machines—has been used in support vector machines^{9,10}, reservoir computing^{11–15}, random mappings^{16–19} and extreme learning machines^{20,21}. The nonlinear mapping performed by the MMF is not the same as in any of the earlier approaches. As we will show, it proves to be effective in transforming the input data space on the SLM to a nearly linearly separable output data space (camera at end of the MM fiber) at high speed and power efficiency.

In the rest of the paper we present numerical and experimental results from our optical computing framework for single-variable linear regression, multivariable linear regression, age prediction from images of faces, audio speech classification and COVID-19 diagnoses from X-ray images tasks. We then discuss how the system scales to a large data size and estimate the power consumption per operation. These studies show that the analog optical computer

¹Optics Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. ²Laboratory of Applied Photonics Devices, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. ✉e-mail: ugur.tegin@epfl.ch

Fig. 1 | Illustration of the neuromorphic computing architectures and the experimental set-up. The experimental set-up consists of an SLM for encoding information onto laser pulses, an MMF for nonlinear projection with spatiotemporal fiber nonlinearities, and an imaging system to decode information. The inset depicts fixed-parameter neural network architectures with similar attributes as the MMF processor, where black and blue connections indicate fixed and adaptable weights, respectively.

based on the MMF is power efficient, versatile and obtains performance comparable with that obtained by digital computers when solving the tasks we investigated (see Supplementary Discussion 10 for detailed comparison).

Results

Experimental studies. Multimode fibers exhibit waveguide properties while allowing a large number of spatial degrees of freedom. Graded-index multimode fibers (GRIN MMFs) in particular have become the subject of considerable interest for telecommunications, imaging and nonlinear optics studies due to their unique properties such as relatively low modal dispersion and periodic self-imaging. With spatiotemporal pulse propagation in GRIN MMFs, nonlinear beam cleaning²², spatiotemporal mode-locking^{23–25} and various nonlinear frequency generation dynamics^{26–30} have been realized in recent years. Moreover, learning and controlling nonlinear optical dynamics in GRIN MMFs was demonstrated by modifying the spatial properties of the intense pump pulse with a SLM or deformable mirror device^{31,32}.

In machine learning studies, a variety of nonlinear transformations of the input data have been investigated to enable learning of complex relations hidden in the data³³. In our case, we make use of the nonlinear mapping that takes place at high light intensities when an input pattern propagates in a MMF as a physical realization of machine learning. The experimental set-up in Fig. 1 is explained in detail in the Methods. In this set-up, information spatially modulated an intense laser pulse with the input data, and the Fourier transform of the spatially modulated beam was focused on the input facet of the optical fiber through a lens. The amount of light coupled to each of the modes of the fiber is given by the inner product between the incident light amplitude and the mode profile. Following propagation, the initial complex modal coefficients evolve according to spatiotemporal linear and nonlinear effects.

The nonlinear transformation of information is achieved by nonlinear energy exchange between the fiber modes. The transformed information at the end of the fiber is imaged onto a camera, and the image was downsampled such that the spatial sampling period is approximately equal to the resolution limit, which can be approximated by $\lambda/2(\text{NA})$, that is, the Abbe diffraction limit, where λ is the central wavelength and NA is numerical aperture. Each pixel of the downsampled image served as an input feature to a linear regression or, equivalently, to a single-layer neural classification algorithm to estimate the identity of the input on the SLM.

The number of modes N in a MMF scales proportionally to the fiber core area, hence the optical power necessary to maintain the same intensity scales linearly with the number of modes. As a fiber with N modes can accommodate an N -dimensional input (as per the law of étendue), the optical power scales with the size of the dimensional input N . The GRIN MMF used in the present study supports 240 modes (counting the polarization degeneracy). In the experiments, learning reached an optimum for a pulse peak power of 3.4 kW for nonlinear optical effects, which corresponds to 4.4 mW average optical power at 125 kHz repetition rate and 10 ps pulse duration.

Learning a nonlinear function. We selected a simple regression problem on a dataset generated with a nonlinear (sinc function) relation. The inputs (x) were randomly generated numbers between $-\pi$ and π , and the corresponding output labels (y) were generated according to the $y = \text{Sin}(\pi x)/(\pi x)$ relation. This dataset is often used as a benchmark in machine learning studies^{20,21}. Each input value (x) was uniquely coded as a two-dimensional pattern, which was recorded on the SLM (see Methods for details). By recording the nonlinearly propagated beam profile of many such input values, a linear regression method was performed on the output data (see Fig. 2a). To measure the effectiveness of the spatiotemporal nonlinear

Fig. 2 | Learning a nonlinear function. Dependence on the pulse propagation regimes in a GRIN MMF, caused by increasing the input optical peak power, is shown. **a**, Measured linear regression results of SOLO outputs for the sinc dataset with indicated labels. Each illustrated sinc example contains 3,000 data points. **b**, An illustration of the propagation difference for linear (low peak power) and nonlinear (high peak power) cases in a GRIN MMF with self-imaging period lengths of ten ($10p$).

propagation and assess the importance of the nonlinearity, we experimented with different pulse peak powers to control the level of nonlinearity. At low peak power (~ 1.14 kW), the nonlinearity is relatively weak and transmission through the fiber is very nearly a linear transformation (except for the square-law at the detector)—performance as a result is poor as the mapping is not linearly separable. Increasing the laser peak power results in nonlinear propagation and performance was optimum at around 3.43 kW laser peak power. For this optimum power level, correct outputs were estimated from unseen test inputs with a r.m.s.e. of 0.0671. Further power escalation gradually drives nonlinear pulse propagation to the Raman beam clean-up regime³⁴, in which the projected beam profiles become virtually unaffected by the input data. In the other experiments reported below, the optical peak power used was 3.43 kW, which corresponds to the optimal peak power of the sinc and COVID-19 diagnosis experiments.

Abalone dataset. We moved to multivariable inference problems and tested our computing method on the abalone dataset³⁵, which consists of various physical features of sea snails that are related to age (for example, the number of rings) and can be used for the prediction of the age of sea snails from eight different parameters. We recorded these eight parameters on the SLM as a 4×2 matrix with proper pixel scaling. The recorded spatial distribution at the distal fiber facet was recorded, flattened and fed into the decision layer to perform linear regression (see Methods). Figure 3a presents the true ages and the corresponding predictions, indicating that the framework learns the ages of the abalone from spatially distributed independent variables with remarkable accuracy (r.m.s.e. of 0.126) compared with the output that takes normalized values between 0 and 1.

Face image dataset. A dataset containing 9,780 images of faces of people of different genders and ethnicities with a long age span (0–116) is used³⁶. The age is first normalized from 0 to 1. The number 1 represents the oldest person (116 years old). The achieved r.m.s.e. for

age prediction is 0.167 normalized years. For the first 1,000 samples, the true ages and predictions are shown in Fig. 3b. Some predictions have negative values, which is impossible; this error is due to the final regression layer. This task is promising as image problems are massive and power-hungry in digital machine learning tools and give rise to convolutional neural network architectures.

Audio digit dataset. We next employed spoken digit classification to challenge the SOLO system. The audio digit classification dataset incorporates recordings of English digits by six distinct people³⁷. Audio recordings are inherently time-varying signals. Following the standard approach, one-dimensional audio signals were converted into two-dimensional representations by generating so-called Mel spectrograms. These spectrograms of audio recordings were provided as inputs into the SLM. The output decision layer classifies the recorded respective fiber output intensity images and 94.5% accuracy over test data is obtained (see Fig. 3c) for a digit categorization task using the frequency-resolved beam profile measurement technique (see Methods).

To demonstrate the versatility of SOLO, we changed the task for the same dataset and aimed to differentiate the speaker from the audio record. As the nonlinear transformation is independent of the task, we only updated the decision layer in SOLO and achieved 95.2% accuracy on test data (as presented in Fig. 3d) using the frequency-resolved beam profile measurement technique (see Methods). The evolution of the loss and accuracy (if applicable) functions for our digital decision layer with SOLO are presented in Fig. 3e–g.

COVID-19 dataset. Encouraged by the performance we obtained with the relatively simple tasks described so far, we tested SOLO with a difficult challenge of current interest by studying COVID-19 diagnoses with a dataset consisting of 3,000 X-ray samples³⁸. Similar to the audio dataset, the X-ray samples are applied to pulses as spatial phase modulation and the corresponding fiber output intensity patterns were recorded. By performing

Fig. 3 | Experimental learning results for regression and classification tasks. a, Measured linear regression results of SOLO outputs for the abalone dataset (3,000 data points are shown) with indicated labels. **b,** Measured linear regression results of SOLO outputs for the human face dataset (3,000 data points are shown) with indicated labels. **c,d,** Confusion matrices for the audio dataset digit (**c**) and speaker (**d**) classification. Color bars show the occurrence of the number of instances and they are unitless. See Methods section for details on confusion matrices. **e,** The evolution of the loss function (mean squared error) for the age prediction dataset. **f,** The evolution of the accuracy and loss functions (categorical cross-entropy) for the digit classification with audio dataset. **g,** The evolution of the accuracy and loss functions (categorical cross-entropy) for the speaker classification with audio dataset.

Fig. 4 | Experimentally tested COVID-19 X-ray dataset. a, A confusion matrix for condition classification. **b,c,** The evolution of the accuracy (**b**) and loss (**c**) functions (binary cross-entropy) for the COVID-19 X-ray dataset. **d,** The measured effect of nonlinear pulse propagation on categorization accuracy. Color bars show the occurrence of the number of instances and they are unitless. See Methods section for details on confusion matrices.

classification in the decision layer, 83.2% accuracy over the unseen test set is achieved (see Fig. 4) with frequency-resolved beam profile measurement technique (see Methods).

To evaluate the robustness of the present optical computing method, we repeated the experiments for the COVID-19 dataset one week after the prior measurements presented in Fig. 4.

Fig. 5 | Illustration of the nonlinear pulse propagation as deep neural network in the experimental set-up. Each elementary step Δz in the MMF is modeled as a cascade of linear and nonlinear network-like operators capturing the phase delay (in blue), the linear mode coupling due to bending and imperfections (in red), and the nonlinear mode coupling (in magenta and yellow). $\hat{D}$ is the diffraction term and $\varphi(x,y)$ is the spatial phase modulation.

Without requiring a calibration, we obtained the similar learning performance level (around 82% accuracy on the test set) for diagnosing COVID-19 from X-ray images (see Supplementary Discussion 6). Furthermore, we performed detailed analysis to determine the stability of the set-up. As an analog system, equipment (laser source, SLM) used in the SOLO experiments inherit noise-like behaviors such as laser-pointing stability or SLM stability. In our tests (see Supplementary Discussion 8), we obtained high stability with a signal to noise ratio of 12.63.

Physical model of SOLO. The nonlinear mapping performed by the MMF can be investigated by the time-dependent (t) beam propagation method (TBPM) involving the fiber mode amplitudes, A_p (where p is the considered mode index; equation (1))³⁹. In an ideal fiber without imperfections and bending, with low power pulse or continuous-wave light, only the phases of the mode coefficients change at different rates—due to modal and chromatic dispersion—without any intermodal power exchange. This behavior is captured by the first term in equation (1). This results in a linear transformation of the field as it propagates through the fiber.

Mode-coupling caused by perturbations due to fiber bending or by impurities (shown by matrix C) also acts as linear mixing (see the second term in equation (1))^{39,40}.

$$\begin{aligned} \frac{\partial A_p}{\partial z} = & \underbrace{i\delta\beta_0^p A_p - \delta\beta_1^p \frac{\partial A_p}{\partial t} - i\frac{\beta_2}{2} \frac{\partial^2 A_p}{\partial t^2}}_{\text{Dispersion}} + \underbrace{i \sum_n C_{p,n} A_n}_{\text{Linear mode coupling}} \\ & + \underbrace{i\gamma \sum_{l,m,n} \eta_{p,l,m,n} A_l A_m A_n^*}_{\text{Nonlinear mode coupling}} \end{aligned} \quad (1)$$

If the peak power of the pulse is high enough to induce nonlinear behavior in the material, nonlinear mode coupling takes place and it results in a nonlinear operation on the information spatially encoded in the intense pulse throughout the fiber (the third term in equation (1)). As illustrated in Fig. 5, by simply iterating equation (1) with a small propagation step, Δz , nonlinear propagation can be obtained for each step. For each propagation step, the fiber modes couple to each other according to the linear coupling coefficients and the nonlinear coupling tensor, indicated by η . This nonlinear operator can be modeled at each propagation step by multiplying each

three-element combination of mode coefficients with the related entry of the nonlinear mode coupling tensor. In equation (1), β_2 is the group velocity dispersion for the central frequency of the pulse; $\delta\beta_0^{(p)}$ ($\delta\beta_1^{(p)}$) is the difference between first (second) Taylor expansion coefficient of the propagation constant for corresponding (p) and the fundamental mode; γ is the Kerr nonlinearity of fused silica fiber; the indices of summations (n, l, m) are the iterating mode numbers to cover all of the possible modal interactions, and i is the imaginary unit. Specific definitions of ($A_p, C_{p,n}, \eta_{p,l,m,n}$) can be found in the ‘physical model of SOLO’ section of the Methods.

Numerical studies of SOLO. We performed TBPM simulations to understand the role of nonlinearity in the MMF and analyze its effects on learning (see Methods). We performed a rescaling of the propagation length and pulse peak power to reduce the computation time but include the required optical nonlinearity (as also explained in the Methods). We numerically studied the learning sinc function (see Methods) as well as the abalone, face image and audio datasets. Note that the numerical simulation is only partial due to computational limitations and the scalar nature of the simulation, the simulated fiber supports 120 spatial modes as opposed to the experiments in which the test fiber supports 240 spatial modes. This difference causes information loss for the nonlinear mapping in the numerical studies.

We simulated nonlinear beam propagation in a GRIN MMF by encoding the COVID-19 dataset onto the optical pulses. We propagated 3,000 X-ray samples numerically, followed by classification of the resulting spatial distribution of the pulses. We achieved 70.8% accuracy (see Fig. 6a–c), which is considerably lower than the experimentally obtained classification accuracy (83.2%). For categorization tasks, our numerical results offered lower performances than our experimental studies. The simplifications may not have captured the complex nonlinear mapping occurring in a longer fiber and lower peak power.

We simulated the COVID-19 dataset with lower peak power levels and shorter fiber lengths to understand the impact of those properties. The numerically obtained COVID-19 diagnosis accuracy of 70.8% decreased to 68.8% and 67.2% when the peak power decreased to half and quarter of the initial power, respectively. Similar results were also obtained for shorter propagation lengths such as by decreasing the fiber length from ten to five self-imaging periods, we could achieve 67.5% diagnosis accuracy. A further decrease of fiber length by two self-imaging periods resulted

Fig. 6 | Numerically tested COVID-19 X-ray dataset. a, A confusion matrix for condition classification. **b, c**, The evolution of the accuracy (**b**) and loss (**c**) functions for COVID-19 X-ray dataset. **d, e**, Simulated effect of pulse peak power (**d**) and fiber length (**e**) on categorization accuracy. Color bars show the occurrence of the number of instances and they are unitless. See Methods section for details on confusion matrices.

in 64.5% diagnosis accuracy in our simulations. The numerically achieved confusion matrices for these studies are shown in Fig. 6.

Discussion

Due to the waveguiding properties of the optical fiber, the light can be fully represented by the complex coefficients of N mutually orthogonal modes (vectors) at the input to the fiber and of the same N modes at the output. The amplitude of each of the output modes is computed optically by linear and nonlinear combinations of all of the input modes as shown in equation 1. Also, as Fig. 5 suggests, this computation is performed many times in the fiber. We thus realize that at least N operations are needed to compute the amplitudes of each output mode, and we have N such modes and thus the scaling is N^2 . With the state-of-the-art SLM technology⁴¹, the computing performance (OPs) of SOLO scales to PetaOPs. (see Supplementary Discussion 11 detailed calculation including the energy efficiency).

The presented optical computing framework can be further improved with an active MMF scheme in which the fiber is mechanically perturbed⁴² or the pump light is also shaped to control spatiotemporal nonlinear propagation. Different cases of adaptive pumping in fiber amplifiers are already demonstrated in the literature^{43,44}. Such an implementation may lead to optically controllable computing with nonlinear fiber optics.

The SOLO framework can be realized with silicon-on-insulator technology, which enables optical functions on integrated circuits and has already resulted in many useful applications⁴⁵. Nonlinear silicon photonics already demonstrated supercontinuum generation through self-phase modulation, light amplification using the Raman effect and matrix convolution operations^{46,47}. By leveraging the existing platform, it is possible to implement the framework that we demonstrated in optical fibers.

In our benchmarks, the proposed optical computing platform performs as powerful as its digital counterparts for different tasks. With better energy efficiency than previous proposals and a path to PetaOPs scalability, SOLO provides a promising path toward super-computer-level optical computation.

Methods

Experimental set-up. As the light source, an yttrium fiber laser (Amplitude Laser Satsuma) that produces 10 ps pulses with a 125 kHz repetition rate was selected. The pulse was centered around 1,033 nm with a width of 10 nm. The linearly polarized Gaussian laser output beam was shaped via a phase-only SLM (Holoeye Pluto-NIRII), an 8 μ m pixel pitch and 60 Hz speed. We illuminated the 600 $\times$ 600 pixels central region of the SLM and all images were scaled to that size to cover the entire beam. A blazed grating was added to the pattern to prevent unmodulated direct current light from entering the fiber. Encoding two-dimensional arrays was relatively easier than encoding a scalar or one-dimensional input. To handle

a scalar input (such as for the sinc experiment), we mapped the scalar value to a two-dimensional array by multiplying the value through a fixed random two-dimensional matrix. This provided unique two-dimensional matrices for every distinct input value. For one-dimensional inputs, we simply converted them into two-dimensions and upsampled them to the illumination pixel range. We used 5 m of a commercial GRIN 50/125 MMF with a NA of 0.2; this fiber allows 120 modes per polarization for the given excitation. The phase-modulated light from SLM was imaged onto the MMF with a lens focal length of 15 mm. The information beam covers the whole MMF core area. The beam-core overlap was checked by imaging the back reflection of the proximal fiber side (not shown in Fig. 1). The distal fiber side was magnified 12.5 times through a 4f imaging system and recorded using a camera with a 5.2 μm pixel pitch. As an alternative method, instead of 4f imaging, frequency-resolved spatial measurement with a dispersive optical element (grating with a 600 line mm⁻¹ period) was used as presented in Fig. 1. Considerable performance increases were observed and reported here for the categorization tasks (audio digit and COVID-19 datasets). We continuously monitored fiber output power after and before the MMF. Various neutral density filters were embedded to avoid camera saturation. The pulse power and width were optimized so that the pulse conserves its temporal unity (no temporal splitting) and maximizes spatial interactions.

Physical model of SOLO. The optical beam at any position of the optical fiber can be decomposed into spatial modes of the fiber.

$$E(\rho, \varphi, \omega) = \sum_l F_l(\rho, \varphi, \omega) e^{i\beta_l(\omega)z} \tilde{A}_l(z, \omega)$$

where, $E(\rho, \varphi, \omega)$ is the electric field of the light, A_l is the envelope of the corresponding mode amplitude along the propagation direction z and F_l is the mode shape.

$$F_{(p,m)}(\rho, \varphi, \omega) = \sqrt{\frac{p!}{\pi(p+|m|)!}} \frac{p^{|m|}}{(w_0/\sqrt{2})^{|m|+1}} e^{-\rho/w_0^2} L_p^{|m|} \left(2 \frac{\rho^2}{w_0^2} \right) e^{im\varphi}$$

shows the solution of the mode shape F_l for graded index fibers having relative index difference of Δ and radius of R , L_p is the generalized Laguerre polynomial³⁹. The values of modes propagation constant $\beta_{p,m}$ were calculated using the following expression (k is the wavenumber),

$$\beta_{p,m}(\omega) = k \sqrt{1 - \frac{2\sqrt{\Delta}}{kR} (2p + |m| + 1)}$$

Propagation of an intense pulse inside an optical fiber can be analyzed following this representation. Equation (1) represents the nonlinear spatiotemporal evolution of each mode. Each mode is coupled to the others through the nonlinearity tensor coefficient,

$$\eta_{p,l,m,n} = \frac{\int dx dy F_p F_l F_m F_n}{\left[\int dx dy F_p \int dx dy F_l \int dx dy F_m \int dx dy F_n \right]^2}$$

which models nonlinear intermodal and intramodal effects, and through a linear coupling coefficient,

$$C_{p,n} = \frac{\omega}{2} \iint dx dy (\tilde{\epsilon}^* - \epsilon) F_p F_n^*$$

which expresses mode coupling due to perturbations to the ideal fiber shape and refractive index distribution, such as bending and impurities. The linear coupling coefficient ($C_{p,n}$) relates perturbation in permittivity (or refractive index) to intermodal model coupling by calculating the overlap integral with the corresponding mode shapes⁴⁰. Similarly, the nonlinearity tensor coefficient ($\eta_{p,l,m,n}$) is computed with the normalized overlap integral of modes. We computed the nonlinear coupling tensor between the modes for our GRIN-50/125 fiber. The tensor has a 120⁴ size and computing all nonlinear terms took two and a half months on a server computer with two Intel Xeon E5-2670 CPUs and 384 GB of RAM. The cross-phase modulation coefficients ($\eta_{p,p,q,q}$) are shown in Supplementary Discussion 1, where the diagonal terms correspond to self-phase modulation ($\eta_{p,p,p,p}$). This demonstrates the richness of the nonlinear interaction that SOLO relies on. Note that four-wave mixing could not be shown on the graph due to its dimensionality.

Numerical simulations. We implemented a GPU-parallelized TBPM,

$$\frac{\partial A}{\partial z} = \frac{i}{2k_0} \left(\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} \right) - i \frac{\beta_2}{2} \frac{\partial^2 A}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3 A}{\partial t^3} - \frac{ik_0 \Delta(x^2 + y^2) A}{R^2} + i\gamma |A|^2 A$$

in Python to simulate sufficiently fast nonlinear pulse propagation in the fiber. Time-dependent beam propagation method simulations often require long computational times due to heavy multidimensional fast Fourier-transform calculations. The launched pulses centered at 1,030 nm with 1 ps duration were numerically propagated for a self-imaging periods distance of 10. In the experiment, the fiber length is 5 m. To reach a manageable computing time for the datasets with 3,000 samples, we performed a rescaling of the propagation length from 5 m to ~5.5 mm. The datasets were divided with a ratio of 0.2 for training (2,400 samples) and validation (600 samples). We increased the pulse peak power to 10 MW to generate nonlinear spatiotemporal evolution in such a short propagation. The time window of simulation is 20 ps with 9.8 fs resolution and the spatial window is set as a 64 × 64 spatial grid. To properly simulate the GRIN MMF's spatial self-imaging, the numerical integration step is set to sample each self-imaging period 16 times. We truncated the parabolic fiber index profile with the super-Gaussian filter to create an absorptive boundary condition around the core. We matched the launched Gaussian beam diameter ($1/e^2$) to fiber core size (50 μm). For our studies, we encoded data into the beam as a multiplied phase information. After propagation, the obtained pulse is time-averaged and converted into normalized intensity images. There were several ways of converting images into one-dimensional representations. For simplicity, we used a flattened version of downsampled images as an output vector. The regression and classification were implemented using Scikit-learn or Tensorflow on Google Colab cloud service which provides an Intel Xeon CPU and a Nvidia Tesla V100 GPU.

We simulated the nonlinear pulse propagation with the same datasets to compare them with our experimental results. We encoded the information as the spatial phase distribution of a pulse in our numerical implementation onto the input beam. Our first numerical simulation was learning the sinc function input-output relation, numerically duplicating the experiment we described above. We used 3,000 randomly generated samples that lie in $[-\pi, \pi]$ to cover the sinc function's characteristic behavior. We fed the generated samples into TBPM by expanding scalar values into a two-dimensional form using a random mask and calculated the nonlinear pulse propagation in the GRIN MMF. The projected intensity distribution at the distal end of the fiber was considered as the nonlinearly transformed information. The linear regression parameters were retrieved from the training data and the overall performance was assessed by the test data. A remarkable learning performance with an r.m.s.e. of 0.0039 for the test data was obtained (see Supplementary Discussion 2). Similar to the sinc function, a decision layer to perform linear regression was employed and we obtained an abalone age prediction with remarkable accuracy (r.m.s.e. of 0.0831). We continued our numerical studies with the face image dataset. By encoding different human face images into the simulated pulse, each person's age on the images was estimated, and an r.m.s.e. of 0.2175 on normalized output values again indicated close correspondence with the experimental studies. We simulated the nonlinear propagation of audio digit pulses for categorization purposes. By taking the fiber output beam shapes as inputs to a single layer classifier gave approximately 68% accuracy. We updated the decision layer and tried to differentiate the speaker from the audio record. In our numerical studies, we obtained 61% accuracy over the unseen test set. The learning results, evolution of loss and accuracy (if applicable) functions for our digital decision layer with experimental results are presented in Supplementary Discussion 3.

Confusion matrices. Each row of the matrix represents the instances in a predicted class and each column represents the instances in an actual (true) class. Each cell shows how many times that class is predicted while the true answer was the label of the corresponding column. Thus, the numbers in the diagonal cells are how many times the correct answers are predicted while the numbers in other cells are false predictions/errors. Color bars show the occurrence of the number of instances and they are unitless. They visualize the distribution number of predictions on the matrix.

Data availability

The datasets containing the raw information for abalone dataset is from (<https://archive.ics.uci.edu/ml/datasets/Abalone>), the face image dataset is from ref.³⁶, audio digit dataset is from (<https://github.com/Jakovovski/free-spoken-digit-dataset>) and COVID-19 dataset is from (<https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>). The recorded experimental and calculated numerical data are available at https://github.com/ugurtegin/Nonlinear_MMf_Network (ref.⁴⁸). Source Data are provided with this paper.

Code availability

The numerical data used in this work and a public version of the codes are available at https://github.com/ugurtegin/Nonlinear_MMf_Network (ref.⁴⁸).

Received: 31 December 2020; Accepted: 15 July 2021;
Published online: 20 August 2021

References

- Cutrona, L., Leith, E. N., Palermo, C. & Porcello, L. Optical data processing and filtering systems. *IRE Trans. Inf. Theory* **6**, 386–400 (1960).
- Cutrona, L. J., Leith, E. N., Porcello, L. J. & Vivian, W. E. On the application of coherent optical processing techniques to synthetic-aperture radar. *Proc. IEEE* **54**, 1026–1032 (1966).
- Miller, D. A. B. et al. Novel hybrid optically bistable switch: the quantum well self-electro-optic effect device. *Appl. Phys. Lett.* **45**, 13–15 (1984).
- Goodman, J. W., Leonberger, F. J., Kung, S. Y. & Athale, R. A. Optical interconnections for VLSI systems. *Proc. IEEE* **72**, 850–866 (1984).
- Feldman, M. R., Esener, S. C., Guest, C. C. & Lee, S. H. Comparison between optical and electrical interconnects based on power and speed considerations. *Appl. Opt.* **27**, 1742–1751 (1988).
- Farhat, N. H., Psaltis, D., Prata, A. & Paek, E. Optical implementation of the Hopfield model. *Appl. Opt.* **24**, 1469–1475 (1985).
- Psaltis, D. et al. Holography in artificial neural networks. *Nature* **343**, 325–330 (1990).
- Wetzstein, G. et al. Inference in artificial intelligence with deep optics and photonics. *Nature* **588**, 39–47 (2020).
- Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
- Noble, W. S. What is a support vector machine? *Nat. Biotechnol.* **24**, 1565–1567 (2006).
- Gallicchio, C. & Micheli, A. Tree echo state networks. *Neurocomputing* **101**, 319–337 (2013).
- Maass, W. In *Computability in Context: Computation and Logic in the Real World* 275–296 (Imperial College Press, 2011).
- Larger, L. et al. Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing. *Opt. Express* **20**, 3241–3249 (2012).
- Brunner, D., Soriano, M. C., Mirasso, C. R. & Fischer, I. Parallel photonic information processing at gigabyte per second data rates using transient states. *Nat. Commun.* **4**, 1–7 (2013).
- Vandoorne, K. et al. Experimental demonstration of reservoir computing on a silicon photonics chip. *Nat. Commun.* **5**, 1–6 (2014).
- Saade, A. et al. Random projections through multiple optical scattering: approximating kernels at the speed of light. *IEEE International Conference on Acoustics, Speech and Signal Processing* 6215–6219 (IEEE, 2016).
- Rafayelyan, M., Dong, J., Tan, Y., Krzakala, F. & Gigan, S. Large-scale optical reservoir computing for spatiotemporal chaotic systems prediction. *Phys. Rev. X* **10**, 041037 (2020).
- Paudel, U., Luengo-Kovac, M., Pilawa, J., Shaw, T. J. & Valley, G. C. Classification of time-domain waveforms using a speckle-based optical reservoir computer. *Opt. Express* **28**, 1225–1237 (2020).
- Sunada, S., Kanno, K. & Uchida, A. Using multidimensional speckle dynamics for high-speed, large-scale, parallel photonic computing. *Opt. Express* **28**, 30349–30361 (2020).
- Huang, G. B., Zhu, Q. Y. & Siew, C. K. Extreme learning machine: theory and applications. *Neurocomputing* **70**, 489–501 (2006).
- Marcucci, G., Pierangeli, D. & Conti, C. Theory of neuromorphic computing by waves: machine learning by rogue waves, dispersive shocks, and solitons. *Phys. Rev. Lett.* **125**, 093901 (2020).
- Krupa, K. et al. Spatial beam self-cleaning in multimode fibres. *Nat. Photon.* **11**, 237–241 (2017).
- Wright, L. G., Christodoulides, D. N. & Wise, F. W. Spatiotemporal mode-locking in multimode fiber lasers. *Science* **358**, 94–97 (2017).
- Tegin, U., Kakkava, E., Rahmani, B., Psaltis, D. & Moser, C. Spatiotemporal self-similar fiber laser. *Optica* **6**, 1412–1415 (2019).
- Tegin, U., Rahmani, B., Kakkava, E., Psaltis, D. & Moser, C. Single-mode output by controlling the spatiotemporal nonlinearities in mode-locked femtosecond multimode fiber lasers. *Adv. Photon.* **2**, 1–8 (2020).
- Wright, L. G., Christodoulides, D. N. & Wise, F. W. Controllable spatiotemporal nonlinear effects in multimode fibres. *Nat. Photon.* **9**, 306–310 (2015).
- Krupa, K. et al. Observation of geometric parametric instability induced by the periodic spatial self-imaging of multimode waves. *Phys. Rev. Lett.* **116**, 183901 (2016).
- Tegin, U. & Ortaç, B. Spatiotemporal instability of femtosecond pulses in graded-index multimode fibers. *IEEE Photon. Technol. Lett.* **29**, 2195–2198 (2017).
- Lopez-Galmiche, G. et al. Visible supercontinuum generation in a graded index multimode fiber pumped at 1064 nm. *Opt. Lett.* **41**, 2553–2556 (2016).
- Tegin, U. & Ortaç, B. Cascaded Raman scattering based high power octave-spanning supercontinuum generation in graded-index multimode fibers. *Sci. Rep.* **8**, 1–7 (2018).
- Tzang, O., Caravaca-Aguirre, A. M., Wagner, K. & Piestun, R. Adaptive wavefront shaping for controlling nonlinear multimode interactions in optical fibres. *Nat. Photonics* **12**, 368–374 (2018).
- Tegin, U. et al. Controlling spatiotemporal nonlinearities in multimode fibers with deep neural networks. *APL Photon.* **5**, 030804 (2020).
- Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019).
- Terry, N. B., Alley, T. G. & Russell, T. H. An explanation of SRS beam cleanup in graded-index fibers and the absence of SRS beam cleanup in step-index fibers. *Opt. Express* **15**, 17509 (2007).
- Abalone Data Set (Marine Resources Division, 1995); <https://archive.ics.uci.edu/ml/datasets/Abalone>
- Zhang, Z., Song, Y., & Qi, H. Age progression/regression by conditional adversarial autoencoder. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 5810–5818 (2017).
- Free-Spoken-Digit-Dataset (GitHub, 2020); <https://github.com/Jakobovski/free-spoken-digit-dataset>
- COVID-19 Radiography Database (Kaggle, 2020); <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>
- Mafi, A. Pulse propagation in a short nonlinear graded-index multimode optical fiber. *J. Lightwave Technol.* **30**, 2803–2811 (2012).
- Snyder, A. W. Coupled-mode theory for optical fibers. *JOSA* **62**, 1267–1277 (1972).
- GAEA-2 10 Megapixel Phase Only LCOS-SLM (Reflective) (Holoeye, 2020); <https://holoeye.com/gaea-4k-phase-only-spatial-light-modulator/>
- Resisi, S., Viernik, Y., Popoff, S. M. & Bromberg, Y. Wavefront shaping in multimode fibers by transmission matrix engineering. *APL Photon.* **5**, 036103 (2020).
- Sperber, T., Billault, V., Dussardier, B., Gigan, S. & Sebbah, P. Gain As configurable disorder-adaptive pumping for control of multimode fiber amplifiers and lasers. Preprint at <https://arxiv.org/abs/2008.04085> (2020).
- Bachelard, N., Gigan, S., Noblin, X. & Sebbah, P. Adaptive pumping for spectral control of random lasers. *Nat. Phys.* **10**, 426–431 (2014).
- Bogaerts, W. et al. Programmable photonic circuits. *Nature* **586**, 207–216 (2020).
- Leuthold, J., Koos, C. & Freude, W. Nonlinear silicon photonics. *Nat. Photon.* **4**, 535–544 (2010).
- Feldmann, J. et al. Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52–58 (2021).
- ugurtegin/Nonlinear_MMf_Network: Final review (Zenodo, 2021); <https://doi.org/10.5281/zenodo.5090719>.

Acknowledgements

We thank N. Ulaş Dinċ and Y. Abu-Mostafa for fruitful discussions. We received no specific funding for this work.

Author contributions

U.T., M.Y. and I.O. performed simulations and experiments, C.M and D.P. supervised and directed the project. All the authors participated in the analysis of the data and the writing process of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43588-021-00112-0>.

Correspondence and requests for materials should be addressed to U.T.

Peer review information *Nature Computational Science* thanks Sylvain Gigan, Hailu Luo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Handling editor: Jie Pan, in collaboration with the *Nature Computational Science* team.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021