Machine Learning and Applications in Ultrafast Photonics

Goëry Genty,[1] Lauri Salmela,[1] John M. Dudley,[2] Daniel Brunner,[2]
Alexey Kokhanovskiy,[3] Sergei Kobtsev,[3] and Sergei K. Turitsyn[3, 4]

[1]*Laboratory of Photonics, Tampere University, FI-33101 Tampere, Finland*

[2]*Institut FEMTO-ST, Université Bourgogne Franche-Comté*
CNRS UMR 6174, 25000 Besançon, France

[3]*Division of Laser Physics and Innovative Technologies,*
Novosibirsk State University, Pirogova str., 2, Novosibirsk, 630090, Russia.

[4]*Aston Institute of Photonic Technologies,*
Aston University, B4 7ET, Birmingham, United Kingdom.

Abstract

Recent years have seen the rapid growth and development of the field of smart photonics, where machine learning algorithms are being matched to optical systems to add new functionalities and to enhance performance. An area where machine learning shows particular potential to accelerate technology is the field of ultrafast photonics – the generation and characterization of light pulses, the study of light-matter interactions on short timescales, and high-speed optical measurements. Our aim here is to highlight a number of specific areas where the promise of machine learning in ultrafast photonics has already been realized, including the design and operation of pulsed lasers, and the characterization and control of ultrafast propagation dynamics. We also consider challenges and future areas of research.

Machine learning is an umbrella term describing the use of statistical techniques and numerical algorithms to carry out tasks without explicit programmed and procedural instructions. Machine learning algorithms are widely used in many areas of engineering and science, with particular strengths in classification, pattern recognition, prediction, system parameter optimization, and the construction of models of complex dynamics from observed data. Machine learning tools have been widely applied in fields such as control systems, speech processing, neuroscience and computer vision [1].

In optics and photonics, early applications of machine learning have mostly been in the form of genetic algorithms for pattern recognition [2], image reconstruction [3], aberration corrections [4], or the design of optical components [5, 6]. More recent work has focused on the analysis of large data sets [7, 8] and on inverse problems where the superior ability of machine learning to classify data, to identify hidden structures and to deal with a large number of degrees of freedom have led to a many results. Particular areas of success include in the design of nanomaterials and structures with specific target properties [9–11], label-free cell classification [12], super resolution microscopy [13, 14], quantum optics [15], and optical communications [16–18].

In addition to applications in the general area of data processing, there is particular potential for machine learning methods to drive the next generation of ultrafast photonic technologies. This is not only because there is increasing demand for adaptive control and self-tuning of ultrafast lasers, but also because many ultrafast phenomena in photonics are nonlinear and multi-dimensional with noise-sensitive dynamics that are extremely challenging to model using conventional methods. While advances in measurement techniques have led to significant progress in experimental studies of such complex dynamics, recent research has shown how machine learning algorithms are providing new ways to identify coherent structures within large sets of noisy data, and can even potentially be applied to determining underlying physical models and governing equations based only on the analysis of complex time series.

Our aim here is to review a number of specific areas where the promise of machine learning in ultrafast photonics has already been realized, and to also consider challenges and future directions of study as well as application where significant impact is expected in the coming years. Before presenting specific details, we first illustrate in Fig. 1 an overview of different machine learning strategies and associated architectures, listing the core concepts, implementation methodologies, and applications where these have been applied in ultrafast photonics.

LASER DESIGN AND SELF-OPTIMIZATION

Self-tuning of ultrafast fibre lasers

Ultrafast lasers are essential tools in many areas of photonics including telecommunications, material processing, and biological imaging [19–23]. They have also played a central role in several Nobel prizes awarded for femtosecond coherent control (1999); the development of the precision frequency comb (2005); and more recently the generation of high-power femtosecond pulses via chirped pulse amplification (2018). Although some ultrafast sources are based on relatively simple designs, the operation of many important laser systems is in fact very complex with dynamic pulse shaping determined by the interplay between a range of nonlinear, dispersive, and dissipative effects [24]. Although this complexity certainly creates challenges in controlling and optimizing the laser emission, it also offers considerable performance advantage not available with simpler systems. A key challenge is then to harness this complexity.

The difficulty in optimizing a particular ultrafast laser arises from the number of degrees of freedom (or control parameters) that need to be balanced to achieve stable operation or reach a specific dynamical regime. Of course, efforts to develop self-optimized or auto-tuned lasers have been made for many years, with the dominant approach being to linearly sweep through a subset of the available parameter space while monitoring the laser output and using a feedback loop to obtain and maintain a desired operating state. While this is a straightforward approach for simpler laser designs with limited parameters, it becomes intractable when the laser operation depends on many degrees of freedom, or when multiple output characteristics need to be optimized simultaneously. Moreover, there is an increasing demand in both research and industrial applications for fully autonomous operation and active realignment in the presence of external perturbations, as well as for the ability to make dynamic changes in pulse characteristics adapted to the target environment (e.g. propagation medium or material). It is for such systems with greatly added complexity that approaches based on machine learning are especially promising and desirable.

An important example here is the widespread fibre laser, where polarization control, pump power, spectral filtering and loss combine to create a wide range of possible operating regimes governed by a rich landscape of nonlinear dynamics [25, 26]. Depending on the exact choice of parameters, the same laser can exhibit very different behaviour: continuous-wave lasing, noise-like pulse generation, Q-switching, mode-locking, multiple pulsing and bound states. It is for this multi-variable optimisation problem that machine learning has recently led to a number of dramatic improvements. The general approach has been to combine an algorithmic feedback loop together with the electronic control of intra-cavity elements varying polarization, pump power, and spectral filtering. Figure 3 shows a generic illustration of machine learning strategies, control elements, and output parameters for optimization of ultrafast fibre lasers. Specifically, Figure 3A illustrates the training phase where control electronics and advanced measurement devices are used to probe the parameter space and map the corresponding operation states, respectively. Collected data are then fed to machine learning algorithms for training. Figure 3B shows the self-tuning regime where the operation state of the laser is characterized in real-time with a simplified measurement system fed into the machine learning algorithm controlling the electronics to lock the system to a desired regime. This is where machine learning is particularly powerful as, once trained, the algorithm allows systematic scanning of the parameter space for optimum operation. Examples of machine learning algorithms that can be used are highlighted in Fig. 2, and general guidelines in applying them are provided in Box 1.

Ultrafast fibre lasers mode-locked by nonlinear polarization evolution (NPE) are particularly complex, because a change in the polarization state affects both spectral and temporal pulse shaping, as well as the gain to loss balance in the cavity due to the intrinsic saturable absorber role played by the polarization-dependent losses. The first studies combining an algorithmic feedback loop with some cavity control parameter were in fact proof-of-concept numerical simulations of an NPE fibre laser, where it was shown that multi-pulsing instability could be reduced via filters optimized with a genetic algorithm [27], and that stochastic changes in environmentally-induced birefringence could be mitigated by applying a singular value decomposition method [28] or using variational autoencoders on the birefringence state map [29, 30]. This modelling was rapidly followed by an experimental implementation using a singular fitness function to identify self-starting regimes in an NPE laser [31]. A number of subsequent experiments for various laser configurations (NPE, ring-cavity, figure-of-eight) have used genetic algorithms to achieve self-tuning and auto-setting in different regimes such as Q-switching, mode-locking, Q-switched mode-locking, or the generation of on-demand pulses with different duration and energies [32–36].

Table I summarizes a selection of results that have been obtained to date (extended from [37]), also providing the characteristics of the particular algorithms used in each case. In most of these studies, the feedback loop typically uses an advanced search or genetic algorithm targeting a desired optimal state based on some particular fitness or objective function as the reference criterion. Although these results are highly promising, genetic algorithms have to be carefully designed due to their sensitivity to the initial choice of population which can lead the fitness function to converge toward a local optimum and be detrimental to multistable dynamics often seen in ultrafast lasers. They also cannot accommodate for long-term dependencies, and the fitness function typically monitors a single parameter limiting the operating regime that can be achieved. Another important drawback of genetic algorithms is their relatively slow convergence time on the scale of minutes or even hours (see Table 1). However, recent developments have shown that one can reduce this time considerably using algorithmic modifications that can mimic human logic, with the possibility to lock the laser to a desired operating state and to recover to this state from perturbation in less than one second [38, 39]. Further improvement in self-tuning speed is likely to require algorithms that also include models of the pulse generating mechanism in order to provide more targeted control. Unfortunately, whilst models based on nonlinear Schrödinger-like equations (NLSE) are generally able to reproduce experimental characteristics qualitatively, quantitative comparison with experiments remains challenging. This is because accurate modelling necessitates the knowledge of a wide range of parameters which are not readily accessible in practice (for example, the random birefringence in the fibre). Ultrafast lasers are also stochastic systems and the impact of noise can generally be only reproduced via computationally intensive Monte-Carlo simulations that require the analysis of a very large amount of data. One can anticipate that the use of machine learning techniques for pattern recognition combined with the latest advances in real-time measurement techniques [40, 41] could lead to better understanding of ultrafast laser dynamics, allowing for the construction of laser systems with improved robustness.

Control of coherent dynamics

In addition to directly controlling laser emission as described above, there is widespread use of extra-cavity shaping technology to modify the characteristics of ultrashort pulses and other light sources used in particular applications. Because such optimization can involve multiple parameters that are interconnected in complex ways, this is an area where machine learning can clearly surpass other forms of manual or partially-automatised control.

For example, pulse compression to a transform-limited duration is essential to femtosecond spectroscopy that uses few-cycle laser pulses to probe physical or chemical interactions. Recently, it was shown how an adaptive neural-network algorithm can control a pulse-shaper and accelerate significantly the compression implementation with a convergence speed 100 times faster than that obtained using more conventional evolutionary algorithms (see Fig. 4A) [42]. Similarly, a neural network was used to determine and optimize the parameters of a pulse shaping system composed of a series of dispersive and nonlinear fibre elements in order to generate arbitrary pulse waveforms (parabolic, triangular or rectangular) of desired duration and chirp [43].

Genetic algorithms can also be used for these purposes, and their application to solve highly nonlinear optimisation problems such as fibre supercontinuum generation has also been very successful [44–47]. Using custom pulse train preparation via an integrated pulse-splitter, a genetic algorithm was used to optimize supercontinuum dynamics to maximize spectral intensity in specific wavelength bands [47] (Fig. 4B). In another study, it was shown how Gaussian-like peaks could be generated at desired wavelengths in a supercontinuum spectrum using a genetic algorithm to tailor the spectral phase of the incident ultrashort pulses [46]. Genetic algorithms have also been applied to the design of fibres with optimized dispersion and nonlinearity coefficient to maximise the bandwidth of coherent supercontinuum in the mid-infrared [44].

Ultrafast characterisation

A central element in the application of machine learning to tune an ultrafast laser is the feedback loop coupling the emitted pulses with the laser cavity parameters. Although some success has been obtained through optimization based on measurements of pulse spectra or temporal autocorrelation functions, ideally a feedback signal based on more complete pulse measurements would be desirable. However, such complete pulse characterization on femtosecond and picosecond timescales generally requires complex optical systems, and the retrieval of the field parameters is an inverse problem which can be particularly time-consuming to solve [48].

Recently, deep neural networks have found applications in solving such inverse problems in areas such as coherent imaging [49, 50], imaging through scattering media [51, 52] or super-resolution [53], and they are now also showing great promise in pulse reconstruction. The first attempt to apply a neural network to reconstruct a short pulse actually dates back to the mid-1990's and the first development of frequency-resolved optical gating (FROG) [54], although this was limited in making strong assumptions about the functional form of the pulse being retrieved. In other work, genetic algorithms have also been successfully applied to FROG trace retrieval [55, 56] but pulse retrieval times still took several minutes. More recently, a convolutional network trained on simulated data was used to reconstruct pulses from experimental FROG traces and was shown to be superior to conventional methods even in the presence of high noise (Fig. 4C) [57]. Additional studies have employed convolutional networks to reconstruct pulses from dispersion scan traces [58], or from multimode fibre nonlinear speckle measurements [59]. Phase recovery for image reconstruction [60–63], X-ray pulse characterisation [64, 65] are also among important emerging and growing areas of applications of machine learning techniques.

COMPLEX DYNAMICS AND TRANSIENT INSTABILITIES

Hidden physics models

The application of machine learning to derive predictive models from sparse or noisy measurements has now penetrated research into the study of the basic properties of physical systems. In particular, a new field of "hidden physics models" has arisen where closed-form mathematical models or nonlinear differential equations governing a physical system [66] are identified automatically by analyzing samples of the dynamical data using "physics-informed neural networks". In some cases, the form of the governing equation(s) may be known or assumed in advance, and the goal is to extract only the unknown coefficients [67]. Alternatively, one can combine a neural network with a compressed sensing-like method to only identify the active terms of the equation(s) from a basis of candidate nonlinear functions [68].

Using these approaches, a number of applications in ultrafast photonics have been demonstrated to analyse pulse propagation dynamics in optical fibre or in fibre lasers associated with the generation of localised and dissipative soliton structures (Fig. 4D) [67]. Model-free approaches in the form of reservoir computing (unlike physics-informed neural networks) have also been implemented to predict coherent dynamics in particular cases of soliton-like propagation (Fig. 4D) [69]. At present, however, such work has been based on numerical data only - the next step in this field is clearly to uncover the governing models from experimental data sets.

Another important area of work involves the study of temporal dependencies observed in nonlinear pulse propagation dynamics, where the temporal and spectral intensity profiles at a specific time instant or propagation length depend on the intensity profiles at earlier times or distance. Recurrent neural networks with internal memory (that are traditionally used for processing and predictions of time-series) are particularly well suited to modelling this type of dynamic behaviour. Indeed very recent results exploiting the memory-capacity of recurrent neural networks show how a recurrent neural network with long short-term memory cell architecture can accurately predict the nonlinear propagation dynamics of short pulses for a wide range of scenarios from higher-order soliton compression (where comparison was made with experiment) to octave-spanning supercontinuum generation [70]. In addition to these studies of single-pass nonlinear propagation dynamics, there is clear potential to use recurrent neural networks in predictions of the complex multi-scale intermittence dynamics also seen in optical fibre lasers [71].

Chaotic systems and instabilities

Chaotic modulation instability in NLSE-like systems is one of the most fundamental examples of instability in optics, with analogs in many other physical systems. Indeed, the study of how incoherent noise can "self-organize" within the NLSE to yield coherent breather structures has attracted wide interest, specifically because of possible links with rogue waves and extreme events [72]. However, the complexity of the measurement techniques needed to directly capture such chaotic breathers on ultrafast timescales has imposed severe constraints on the dynamical regimes that can be explored in experiments [73, 74].

Machine learning has been used to address this problem directly by training a neural network to determine the temporal characteristics of a chaotic field based only on the spectral intensity characteristics (which are easier to measure). Using numerical data generated from NLSE simulations, a neural network was used to construct a nonlinear transfer function that maps noisy broadband spectra to the local intensity maximum of the chaotic temporal field (see Fig. 4E). This function was then applied to experimental data measured using a high dynamic range real-time spectrometer [75]. A similar approach was recently used to determine the peak power, duration, and temporal delay of extreme rogue solitons in noisy supercontinuum generation [76]. Also analyzing chaotic data from modulation instability, unsupervised clustering analysis using the k-mean algorithm was shown to successfully sort intensity spectra into sub-classes associated in the time-domain with specific solutions of the NLSE related to analytic soliton structures [75].

The application of machine learning techniques has been extended to even more complex systems such as those observed in transient laser behaviour and extreme events [77]. Specifically, using the knowledge of previous pulses in a chaotic time series from an optically injected semiconductor laser operating, machine learning methods (nearest neighbors, support vector machine, feed-forward neural networks, reservoir computing) were analyzed for their ability to predict the intensity of upcoming pulses emitted from the laser [77, 78]. Although this work was numerical, it clearly shows the potential of such prediction in experiment. Attempts have also been made to model highly incoherent system evolution including multidimensional spatiotemporal systems [79] but the predictions in this case tend to diverge over longer distances [80].

Multidimensional systems

A major benefit of neural networks is their ability to efficiently analyze the properties of multi-dimensional systems. This can be particularly useful in multimode fibre systems where spatio-temporal coupling increases dramatically the parameter space and complexity of nonlinear propagation dynamics. The potential of machine learning in this case was recently demonstrated with experiments tailoring supercontinuum generation in a graded index fibre through control of the injected spatial beam profile via a neural-network driven spatial light modulator [81].

Extension to spatial control for enhanced near-field interactions was also shown by combining a neural network with a genetic algorithm to optimise spectral-phase shaping of an incident field to achieve second harmonic generation hotspot switching in plasmonic nanoantennas [82]. In this latter work, the genetic algorithm was added to generate a wide range of nanoantenna designs to be fed into the neural network.

OUTLOOK AND CHALLENGES

Ultrafast photonics systems are generally very complex, often nonlinear, and with dynamics extremely sensitive to both their internal parameters and external perturbations. The design and optimization of these systems have been typically based on physical models, numerical simulations, and trial-and-error approaches. With the increased complexity of these systems, driven by the demand for high stability, robustness against disturbances, tunability and adaptive control, these approaches are now starting to reach their limits such that future major advances will require new methodologies that can analyse the systems characteristics at a global level. One may therefore anticipate that machine learning techniques able to discover hidden features and independently adapt as they are exposed to new data, are likely to play a central role in the next generation of ultrafast systems and applications. There are of course many ways machine learning techniques can be exploited, and we discuss below some possible future direction of research and challenges to overcome.

Ultrafast fibre lasers are dynamical systems operating in regimes determined by dispersion, nonlinearity, gain, losses, and saturation effects. Optimization, breakthrough performance, high stability against perturbations, and automatic-tuning requires in-depth understanding of the full system parameter space, which can be achieved by combining accurate real-time characterization and advanced data analysis. Machine learning-based approaches have the potential to reduce the complexity and number of measurement devices typically required. They could further allow for converting results of measurements into a higher-dimensional space where the separation of the role played by the different cavity elements is more apparent, aiding the construction of universal models. Machine learning may also yield significant developments in full and high-speed characterization of short pulses or complex fields arising from highly nonlinear dynamics. Adaptive optics and coherent control typically rely on ultrafast laser systems where the spatial, temporal and spectral properties of the laser beam are central to optimum performance in e.g. metrology [83], spectroscopy [84, 85], energy harvesting [86] or astronomy [87]. By enabling more systematic strategies rather than heuristic approaches (e.g. in the optimization of multidimensional systems including beam shaping and space-time focusing in multimode fibers [88–90]), machine learning could enable unprecedented level of control in those applications. Another important area where we expect machine learning to lead to significant progress is the discovery of models using data-driven strategy, allowing for finding governing mathematical equations of complex optical phenomena or photonics systems. It is even conceivable that in the future ultrafast fibre lasers could become testbeds for the physics discovered from machine learning.

To date, the majority of machine learning applications to ultrafast photonics have been based on genetic algorithms or feed forward architectures. While these implementations have undoubtedly led to remarkable and pioneering results, there are still important approaches that have yet to be fully exploited. Indeed, it is likely that realising the full potential of machine learning will necessitate the combination of several strategies that have so far been used only separately. For example, recurrent networks based on long short-term memory cells, gated recurrent units, or reservoir computing that possess internal memory can be used to model dynamical systems consisting of time series of different states. These approaches could enable significant progress in understanding and optimizing nonlinear systems, allowing identification of long-term dependencies and internal dynamics in ultrafast lasers, or the prediction of complex evolution maps associated with the propagation of short pulses in nonlinear media and related instabilities. Also, the capabilities of unsupervised learning to draw inferences and reveal hidden internal structures from data sets without labelled responses could be of significant interest in problems where dimensionality reduction is key. These include e.g. multimodal systems or noise-sensitive dynamics where specific regimes can be divided into a number of different clusters associated with measurable parameter(s). Moreover, approaches employed for the design of nanophotonic components in the form of machine learning combined with the adjoint method [91] could be a powerful tool for the inverse design of ultrafast photonics systems. The concept of generative adversarial networks [92] where two distinct networks are optimized in the backpropagation operation [93] is another promising avenue to explore in ultrafast photonics.

There are of course important challenges ahead. When using recurrent network to analyze and predict dynamics, proper sampling along the evolution dimension (time or distance) is essential to extract and reproduce the long-term evolution structure. Memory limitations can then become an issue especially in the context of lasers where it takes usually many cavity round trips for a regime to stabilize. Unsupervised learning analysis divides the data into subsets with similarities, but crucial information on the criterion used to perform the division, or on what the similarities actually are within the clusters is lacking. This means that in order to fully exploit the power of unsupervised learning, further human investigation is generally needed to establish the link between the clusters and specific parameters of the system analysed. This can be a limiting factor, especially for the case of noise-sensitive systems where tiny variations can result in dramatically different evolution patterns.

The use of machine learning algorithms for real-time processing of photonic systems that can produce data in excess of billions of bits per second requires the ability to manage high data volumes, as well as a hardware framework capable of dealing with ultrafast processing rates. In order to reduce the large volume of data, one could use the approach of spike-based neural networks that can reconstruct features of spatio-temporal states based on a fraction of that regime information. Inspired by the human brain that strongly compresses the information received from the eye [94], spike-based neural networks use a specific set of rules such as spike time-dependent plasticity leading to self-organization of the network's topology and allowing to identify possible correlations in the input data. When combined with lateral inhibition (a spike-based form of a winner take all topology), spiked-based neural networks can self-configure to perform a cluster analysis with performance similar to that achieved with a k-mean algorithm [95]. Efforts to develop a hardware framework allowing for high-speed processing and optimization on short time scales have already been made, and several all-optical network architectures have been proposed based e.g. on multiple layers of diffractive surfaces where each point on a given layer acts as a node [96], or based on optical matrix multiplication using a cascaded array of Mach–Zehnder interferometers integrated into a silicon photonic circuit [97]. Another promising approach could be to combine all-optical field-programmable gate arrays and fully parallel photonic neural network hardware. Of course, one important constraint to the development of all-optical neural net- works that needs to be carefully studied is the tolerance to photonic component fabrication imperfections [98].

In the past few years, there have been remarkable developments enabled by the use of machine learning techniques, and an active field of machine-learning ultrafast photonics has now been established. As research continues to progress both in the development of machine learning algorithms and ultrafast photonics technologies, we can expect even more fruitful interactions with increased influence of the former in the physical understanding, design, optimization, and operation of the latter.

ACKNOWLEDGMENTS

GG acknowledges the Academy of Finland (298463, 318082, Flagship PREIN 320165). LS acknowledges the Faculty of Engineering and Natural Sciences graduate school of Tampere University. JD acknowledges the French Agence Nationale de la Recherche (ANR-15-IDEX-0003, ANR-17-EURE-0002). DB acknowledges funding from the Volkswagen Foundation and from the French Agence Nationale de la Recherche (ANR-19-CE24-0006-02). The work of SKT and AK was supported by the Russian Science Foundation (grant No. 17-72-30006), SKT acknowledge support of the EPSRC project TRANSNET. The work of SK was supported by the Russian Foundation for Basic Research grant no. 18-29-20025.

CORRESPONDENCE

All correspondence should be addressed to Goëry Genty.

References

[1] Jordan, M.I. & Mitchell, T.M. Machine learning: trends, perspectives, and prospects. *Science* **349**, 255–260 (2015).

[2] Mahlab, U., Shamir, J. & Caulfield, H.J. Genetic algorithm for optical pattern recognition. *Opt. Lett.* **16**, 648–650 (1991).

[3] Kihm, K.D. & Lyons, D.P. Optical tomography using a genetic algorithm. *Opt. Lett.* **21**, 1327–1329 (1996).

[4] Albert, O., Sherman, L., Mourou, G., Norris, T.B. & Vdovin, G. Smart microscope: an adaptive optics learning system for aberration correction in multiphoton confocal microscopy. *Opt. Lett.* **25**, 52–54 (2000).

[5] Eisenhammer, T., Lazarov, M., Leutbecher, M., Schöffel, U. & Sizmann, R. Optimization of interference filters with genetic algorithms applied to silver-based heat mirrors. *Appl. Optics* **32**, 6310–6315 (1993).

[6] Martin, S., Rivory, J. & Schoenauer, M. Synthesis of optical multilayer systems using genetic algorithms. *Appl. Optics* **34**, 2247–2254 (1995).

[7] Zibar, D., Wymeersch, H. & Lyubomirsky, I. Machine learning under the spotlight. *Nat. Photonics* **11**, 749–751 (2017).

[8] Zhou, J., Huang, B., Yan, Z. & Bünzli, J.-C.G. Emerging role of machine learning in light-matter interaction. *Light Sci. Appl.* **8**, 84 (2019).

[9] Nadell, C.C., Huang, B., Malof, J.M. & Padilla, W.J. Deep learning for accelerated all-dielectric metasurface design. *Opt. Express* **27**, 27523–27535 (2019).

[10] Malkiel, I. *et al.* Plasmonic nanostructure design and characterization via deep learning. *Light Sci. Appl.* **7**, 60 (2018).

[11] Hegde, R.S. Deep learning: a new tool for photonic nanostructure design. *Nanoscale Adv.* **2**, 1007–1023 (2020).

[12] Chen, C.L. *et al.* Deep learning in label-free cell classification. *Sci. Rep.* **6**, 21471 (2016).

[13] Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468 (2018).

[14] Durand, A. *et al.* A machine learning approach for online automated optimization of super-resolution optical microscopy. *Nat. Commun.* **9**, 5247 (2018).

[15] Palmieri, A.M. *et al.* Experimental neural network enhanced quantum tomography. *npj*

Quantum Inf. **6**, 20 (2020).

[16] Zibar, D., Piels, M., Jones, R. & Schäeffer, C.G. Machine learning techniques in optical communication. *J. Lightwave Technol.* **34**, 1442–1452 (2016).

[17] Musumeci, F. *et al.* An overview on application of machine learning techniques in optical networks. *IEEE Commun. Surv. Tut.* **21**, 1383–1408 (2019).

[18] Lugnan, A. *et al.* Photonic neuromorphic information processing and reservoir computing. *APL Photonics* **5**, 020901 (2020).

[19] Knox, W.H. Ultrafast technology in telecommunications. *IEEE J. Sel. Top. in Quant.* **6**, 1273–1278 (2000).

[20] Sibbett, W., Lagatsky, A.A. & Brown, C.T.A. The development and application of femtosecond laser systems. *Opt. Express* **20**, 6989–7001 (2012).

[21] Sugioka, K. & Cheng, Y. Ultrafast lasers–reliable tools for advanced materials processing. *Light Sci. Appl.* **3**, e149 (2014).

[22] Fermann, M.E., Galvanauskas, A. & Sucha, G. *Ultrafast lasers: technology and applications*, vol. 80 (CRC Press, 2002).

[23] Xu, C. & Wise, F.W. Recent advances in fibre lasers for nonlinear microscopy. *Nat. Photonics* **7**, 875–882 (2013).

[24] Grelu, P. & Akhmediev, N. Dissipative solitons for mode-locked lasers. *Nat. Photonics* **6**, 84–92 (2012).

[25] Richardson, D.J., Nilsson, J. & Clarkson, W.A. High power fiber lasers: current status and future perspectives. *J. Opt. Soc. Am. B* **27**, B63–B92 (2010).

[26] Fermann, M.E. & Hartl, I. Ultrafast fibre lasers. *Nat. Photonics* **7**, 868–874 (2013).

[27] Fu, X. & Kutz, N.J. High-energy mode-locked fiber lasers using multiple transmission filters and a genetic algorithm. *Opt. Express* **21**, 6526–6537 (2013).

[28] Fu, X., Brunton, S.L. & Kutz, J.N. Classification of birefringence in mode-locked fiber lasers using machine learning and sparse representation. *Opt. Express* **22**, 8585–8597 (2014).

[29] Kutz, J.N. & Brunton, S.L. Intelligent systems for stabilizing mode-locked lasers and frequency combs: machine learning and equation-free control paradigms for self-tuning optics. *Nanophotonics* **4**, 459–471 (2015).

[30] Baumeister, T., Brunton, S.L. & Kutz, J.N. Deep learning and model predictive control for

self-tuning mode-locked lasers. *J. Opt. Soc. Am. B* **35**, 617–626 (2018).

[31] Andral, U. *et al.* Fiber laser mode locked through an evolutionary algorithm. *Optica* **2**, 275–278 (2015).

[32] Andral, U. *et al.* Toward an autosetting mode-locked fiber laser cavity. *J. Opt. Soc. Am. B* **33**, 825–833 (2016).

[33] Woodward, R. & Kelleher, E. Towards smart lasers: self-optimisation of an ultrafast pulse source using a genetic algorithm. *Sci. Rep.* **6**, 37616 (2016).

[34] Woodward, R. & Kelleher, E. Genetic algorithm-based control of birefringent filtering for self-tuning, self-pulsing fiber lasers. *Opt. Lett.* **42**, 2952–2955 (2017).

[35] Winters, D.G., Kirchner, M.S., Backus, S.J. & Kapteyn, H.C. Electronic initiation and optimization of nonlinear polarization evolution mode-locking in a fiber laser. *Opt. Express* **25**, 33216–33225 (2017).

[36] Kokhanovskiy, A., Ivanenko, A., Kobtsev, S., Smirnov, S. & Turitsyn, S. Machine learning methods for control of fibre lasers with double gain nonlinear loop mirror. *Sci. Rep.* **9**, 2916 (2019).

[37] Meng, F. & Dudley, J.M. Towards a self-driving ultrafast fiber laser. *Light Sci. Appl. (News & Views)* **9**, 26 (2020).

[38] Pu, G., Yi, L., Zhang, L. & Hu, W. Intelligent programmable mode-locked fiber laser with a human-like algorithm. *Optica* **6**, 362–369 (2019).

[39] Pu, G., Yi, L., Zhang, L. & Hu, W. Genetic algorithm-based fast real-time automatic mode-locked fiber laser. *IEEE Photonics Technol. Lett.* **32**, 7–10 (2020).

[40] Kokhanovskiy, A. *et al.* Machine learning-based pulse characterization in figure-eight mode- locked lasers. *Opt. Lett.* **44**, 3410–3413 (2019).

[41] Pu, G. *et al.* Intelligent control of mode-locked femtosecond pulses by time-stretch-assisted real-time spectral analysis. *Light Sci. Appl.* **9**, 13 (2020).

[42] Farfan, C.A., Epstein, J. & Turner, D.B. Femtosecond pulse compression using a neural-network algorithm. *Opt. Lett.* **43**, 5166–5169 (2018).

[43] Finot, C., Gukov, I., Hammani, K. & Boscolo, S. Nonlinear sculpturing of optical pulses with normally dispersive fiber-based devices. *Opt. Fiber Technol.* **45**, 306–312 (2018).

[44] Zhang, W.Q., Afshar, S. & Monro, T.M. A genetic algorithm based approach to fiber design for high coherence and large bandwidth supercontinuum generation. *Opt. Express*

17, 19311–19327 (2009).

[45] Arteaga-Sierra, F.R. *et al.* Supercontinuum optimization for dual-soliton based light sources using genetic algorithms in a grid platform. *Opt. Express* **22**, 23686–23693 (2014).

[46] Michaeli, L. & Bahabad, A. Genetic algorithm driven spectral shaping of supercontinuum radiation in a photonic crystal fiber. *J. Opt.* **20**, 055501 (2018).

[47] Wetzel, B. *et al.* Customizing supercontinuum generation via on-chip adaptive temporal pulse-splitting. *Nat. Commun.* **9**, 4884 (2018).

[48] Ryczkowski, P. *et al.* Real-time full-field characterization of transient dissipative soliton dynamics in a mode-locked laser. *Nat. Photonics* **12**, 221–227 (2018).

[49] Kamilov, U.S. *et al.* Learning approach to optical tomography. *Optica* **2**, 517–522 (2015).

[50] Rivenson, Y., Wu, Y. & Ozcan, A. Deep learning in holography and coherent imaging. *Light Sci. Appl.* **8**, 85 (2019).

[51] Borhani, N., Kakkava, E., Moser, C. & Psaltis, D. Learning to see through multimode fibers. *Optica* **5**, 960–966 (2018).

[52] Li, Y., Xue, Y. & Tian, L. Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media. *Optica* **5**, 1181–1190 (2018).

[53] Liu, T. *et al.* Deep learning-based super-resolution in coherent imaging systems. *Sci. Rep.* **9**, 3926 (2019).

[54] Krumbügel, M.A. *et al.* Direct ultrashort-pulse intensity and phase retrieval by frequency-resolved optical gating and a computational neural network. *Opt. Lett.* **21**, 143–145 (1996).

[55] Nicholson, J., Omenetto, F., Funk, D.J. & Taylor, A. Evolving FROGS: phase retrieval from frequency-resolved optical gating measurements by use of genetic algorithms. *Opt. Lett.* **24**, 490–492 (1999).

[56] Shu, S.F. Evolving ultrafast laser information by a learning genetic algorithm combined with a knowledge base. *IEEE Photonics Technol. Lett.* **18**, 379–381 (2006).

[57] Zahavy, T. *et al.* Deep learning reconstruction of ultrashort pulses. *Optica* **5**, 666–673 (2018).

[58] Kleinert, S., Tajalli, A., Nagy, T. & Morgner, U. Rapid phase retrieval of ultrashort pulses from dispersion scan traces using deep neural networks. *Opt. Lett.* **44**, 979–982 (2019).

[59] Xiong, W. *et al.* Deep learning of ultrafast pulses with a multimode fiber. *arXiv preprint arXiv:1911.00649* (2019).

[60] Rivenson, Y., Zhang, Y., Günaydın, H., Teng, D. & Ozcan, A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light Sci. Appl.* **7**, 17141 (2018).

[61] Sinha, A., Lee, J., Li, S. & Barbastathis, G. Lensless computational imaging through deep learning. *Optica* **4**, 1117–1125 (2017).

[62] Wu, Y. *et al.* Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery. *Optica* **5**, 704–710 (2018).

[63] Goy, A., Arthur, K., Li, S. & Barbastathis, G. Low photon count phase retrieval using deep learning. *Phys. Rev. Lett.* **121**, 243902 (2018).

[64] Sanchez-Gonzalez, A. *et al.* Accurate prediction of x-ray pulse properties from a free-electron laser using machine learning. *Nat. Commun.* **8**, 15461 (2017).

[65] White, J. & Chang, Z. Attosecond streaking phase retrieval with neural network. *Opt. Express* **27**, 4799–4807 (2019).

[66] Raissi, M. Deep hidden physics models: deep learning of nonlinear partial differential equations. *J. Mach. Learn. Res.* **19**, 932–955 (2018).

[67] Raissi, M., Perdikaris, P. & Karniadakis, G.E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).

[68] Brunton, S.L., Proctor, J.L. & Kutz, J.N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *P. Natl. Acad. Sci.* **113**, 3932–3937 (2016).

[69] Jiang, J. & Lai, Y.-C. Model-free prediction of spatiotemporal dynamical systems with recurrent neural networks: role of network spectral radius. *Phys. Rev. Res.* **1**, 033056 (2019).

[70] Salmela, L. *et al.* Predicting ultrafast nonlinear dynamics in fibre optics with a recurrent neural network. *arXiv preprint arXiv:2004.14126* (2020).

[71] Lapre, C. *et al.* Real-time characterization of spectral instabilities in a mode-locked fibre laser exhibiting soliton-similariton dynamics. *Sci. Rep.* **9**, 13950 (2019).

[72] Dudley, J.M., Genty, G., Mussot, A., Chabchoub, A. & Dias, F. Rogue waves and analogies in optics and oceanography. *Nat. Rev. Phys.* **1**, 675–689 (2019).

[73] Närhi, M. *et al.* Real-time measurements of spontaneous breathers and rogue wave events

in optical fibre modulation instability. *Nat. Commun.* **7**, 13675 (2016).

[74] Tikan, A., Bielawski, S., Szwaj, C., Randoux, S. & Suret, P. Single-shot measurement of phase and amplitude by using a heterodyne time-lens system and ultrafast digital time-holography. *Nat. Photonics* **12**, 228–234 (2018).

[75] Närhi, M. *et al.* Machine learning analysis of extreme events in optical fibre modulation instability. *Nat. Commun.* **9**, 4923 (2018).

[76] Salmela, L., Lapre, C., Dudley, J.M. & Genty, G. Machine learning analysis of rogue solitons in supercontinuum generation. *Sci. Rep.* **10**, 9596 (2020).

[77] Amil, P., Soriano, M.C. & Masoller, C. Machine learning algorithms for predicting the amplitude of chaotic laser pulses. *Chaos* **29**, 113111 (2019).

[78] Cunillera, A., Soriano, M.C. & Fischer, I. Cross-predicting the dynamics of an optically injected single-mode semiconductor laser using reservoir computing. *Chaos* **29**, 113113 (2019).

[79] Vlachas, P. *et al.* Backpropagation algorithms and reservoir computing in recurrent neural networks for the forecasting of complex spatiotemporal dynamics. *Neural Networks* **126**, 191–217 (2020).

[80] Pathak, J., Hunt, B., Girvan, M., Lu, Z. & Ott, E. Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. *Phys. Rev. Lett.* **120**, 024102 (2018).

[81] Teğin, U. *et al.* Controlling spatiotemporal nonlinearities in multimode fibers with deep neural networks. *APL Photonics* **5**, 030804 (2020).

[82] Comin, A. & Hartschuh, A. Efficient optimization of SHG hotspot switching in plasmonic nanoantennas using phase-shaped laser pulses controlled by neural networks. *Opt. Express* **26**, 33678–33686 (2018).

[83] Diddams, S.A. The evolving optical frequency comb. *J. Opt. Soc. Am. B* **27**, B51–B62 (2010).

[84] Assion, A. *et al.* Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses. *Science* **282**, 919–922 (1998).

[85] Bartels, R. *et al.* Shaped-pulse optimization of coherent emission of high-harmonic soft X-rays. *Nature* **406**, 164–166 (2000).

[86] Herek, J.L., Wohlleben, W., Cogdell, R.J., Zeidler, D. & Motzkus, M. Quantum control of

energy flow in light harvesting. *Nature* **417**, 533–535 (2002).

[87] Davies, R. & Kasper, M. Adaptive optics for astronomy. *Annu. Rev. Astron. Astrophys.* **50**, 305–351 (2012).

[88] Florentin, R. *et al.* Shaping the light amplified in a multimode fiber. *Light Sci. Appl.* **6**, e16208 (2017).

[89] Florentin, R., Kermene, V., Desfarges-Berthelemot, A. & Barthelemy, A. Space-time adaptive control of femtosecond pulses amplified in a multimode fiber. *Opt. Express* **26**, 10682– 10690 (2018).

[90] Liu, B. & Weiner, A.M. Space–time focusing in a highly multimode fiber via optical pulse shaping. *Opt. Lett.* **43**, 4675–4678 (2018).

[91] Hughes, T.W., Minkov, M., Williamson, I.A.D. & Fan, S. Adjoint method and inverse design for nonlinear nanophotonic devices. *ACS Photonics* **5**, 4781–4787 (2018).

[92] Goodfellow, I. J. *et al.* Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **3**, 2672–2680 (2014).

[93] Subramaniam, A., Wong, M.L., Borker, R.D., Nimmagadda, S. & Lele, S.K. Turbulence enrichment using physics-informed generative adversarial networks. *arXiv preprint arXiv:2003.01907* (2020).

[94] Van Rullen, R. & Thorpe, S.J. Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Comput.* **13**, 1255–1283 (2001).

[95] Diamond, A., Schmuker, M. & Nowotny, T. An unsupervised neuromorphic clustering algorithm. *Biol. Cybern.* **113**, 423–437 (2019).

[96] Lin, X. *et al.* All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).

[97] Shen, Y. *et al.* Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017).

[98] Fang, M.Y.-S., Manipatruni, S., Wierzynski, C., Khosrowshahi, A. & DeWeese, M.R. Design of optical neural networks with component imprecisions. *Opt. Express* **27**, 14009–14029 (2019).

[99] Young, S.R., Rose, D.C., Karnowski, T.P., Lim, S.-H. & Patton, R.M. Optimizing deep learning hyper-parameters through an evolutionary algorithm. In *Proceedings of the Workshop on Machine Learning in High-Performance Computing Environments*, 1–5

(2015).

[100] Penkovsky, B., Larger, L. & Brunner, D. Efficient design of hardware-enabled reservoir computing in FPGAs. *J. Appl. Phys.* **124**, 162101 (2018).

[101] Klein, A., Falkner, S., Bartels, S., Hennig, P. & Hutter, F. Fast Bayesian hyperparameter optimization on large datasets. *Electron. J. Stat.* **11**, 4945–4968 (2017).

[102] Antonik, P., Marsal, N., Brunner, D. & Rontani, D. Bayesian optimisation of large-scale photonic reservoir computers. *arXiv preprint arXiv:2004.02535* (2020).

[103] Meng, F., Lapre, C., Billet, C., Genty, G. & Dudley, J.M. Instabilities in a dissipative soliton-similariton laser using a scalar iterative map. *Opt. Lett.* **45**, 1232–1235 (2020).

BOX 1. General considerations when applying machine learning models

Choosing an architecture and associated parameters Neural networks are universal function approximators whose performance significantly depends on their hyperparameters (variables that determines the network structure and training). Selecting the optimum architecture (Figs. 1-2) and tuning the hyperparameters often involves significant heuristics, exhaustive scans, trial and error, and leveraged optimization tools (genetic algorithms or Bayesian methods). Nevertheless, one may consider the following guidelines to select an appropriate architecture and hyperparameters: a feedforward neural network is a good choice if the map from input to output lacks temporal context. This is typically the case when one considers input-output mappings of "single-pass" systems such as pulses undergoing nonlinear propagation, where fluctuations are expected to be independent and uncorrelated, and also for particular classes of similarly (partially) uncorrelated instabilities in Q-switched lasers. If data contains structure along a particular input dimension (e.g. space, time or wavelength), architectures including filters such as convolutional neural networks are better candidates; one may employ fully connected topologies for input data apparently lacking such features. If the output is expected to depend on current and past input data, recurrent topologies (long short-term memory, gated recurrent units, or reservoir computing) should be used.

Accuracy generally increases with the number of hidden layers or nodes. The number of layers, nodes and training epochs can be increased until the validation error starts increasing (even if the training error still decreases). Note that too many nodes can lead to overfitting and reduce generalization (the ability of a trained model to adapt accurately to data outside the initial training data set). Continuously reducing the number of nodes for deeper layers is a common strategy to improve generalization, and 2 to 3 hidden layers comprising 50 to 1000 nodes appear sufficient for most tasks in ultrafast photonics. A neural network's inference quality is quantified by a cost function such as mean squared or root mean squared error. The root mean squared error penalizes small divergences more heavily and can be employed when fast and accurate convergence is essential. Network weights are typically initialized randomly, and popular activation functions are the rectified linear unit and the sigmoid nonlinearity. The rectified linear unit is computationally less expensive and avoids vanishing gradients, while the sigmoid's upper limit makes blowing-up solutions less likely.

Selecting training data There is generally no one-size-fits-all criterion to determine the volume of training data needed for a specific network and task. Where possible, one can be guided by available examples of comparable problems, and more generally, an initial guess can be obtained by considering the number of classes (output neurons), relevant input features (e.g. optical modes), and parameters of the underlying model. One can then continuously increase the volume of training data until the validation error stagnates. The training data should be representative of the system's possible states, and therefore sample uniformly the system's phase space. This can be challenging, especially for ultrafast nonlinear systems which may rarely visit specific outlier regions (so-called skewed data-set), and can lead to degraded performance in testing. Feeding representative data sets is also not always possible during experiments, and data augmentation via simulation is an alternative approach. It is also important to normalize training data to the 'useful' range of the neurons' nonlinear response (around unity) so as to prevent the network operating in the linear or saturated regime.

Avoiding overfitting Unlike in genetic algorithms, overfitting can occur in neural networks, typically when the testing error is large compared to the training error. The risk of overfitting may be reduced using the following strategies: simplification to reduce the network complexity; data augmentation by increasing the fraction of noisy data during training; cross-validation where division of data into training and testing sets is varied during training; early stopping where training is stopped when the testing error starts increasing; regularization by including penalties in the system's loss function; drop-out by randomly removing individual connections during training.

Robustness and transfer learning Ultrafast photonic systems are generally sensitive to their environment Enabling stable and robust operation is another key objective for machine learning. Performance degradation upon a change of environmental conditions will mostly depend on the parameter space and regimes explored during training and testing. It is therefore important to include training data that incorporates possible environmental variations (see also Selecting Training Data). Using unsupervised learning to determine the dynamic relation between external conditions and system output is another approach.

A related question is “transfer learning”, or how a neural network architecture optimized for a particular system can be `transferred’ to a different yet related problem. In particular, the output of an ultrafast system can be divided into different regimes depending on the system parameters. This is particularly true for mode-locked laser pulses which typically correspond to fundamental solitons, dissipative solitons, or periodic breathers depending on the laser dispersion, nonlinearity, gain, loss, and filtering. Transfer learning may then use training data generated with simplified mathematical models or experiments with reduced complexity. In fact, transfer learning is in itself an important topic of machine learning research and from that point of view ultrafast photonic devices could be ideal testbeds for investigating transfer learning problems in general.

FIGURE CAPTIONS

FIG. 1. Overview of main machine learning concepts and implementations that can be used in ultrafast photonics. The figure illustrates the core concepts and corresponding implementation methodologies as delimited by the coloured arcs, and links these to particular applications where these have been applied in ultrafast photonics. There are also other concepts including semi-supervised learning and reinforcement learning which use some of the implementations mentioned in the figure, but these have yet to be exploited in an ultrafast context. Of course, we also stress that all these methods have been used in many other fields of science in addition to the ones shown here.

FIG. 2. Widespread and promising machine learning architectures for ultrafast photonics. **A**: Genetic algorithm. **B**: Feed-forward neural network. **C**: Convolutional neural network. **D**: Unsupervised learning. **E**: Recurrent neural network. **F**: Reservoir computing. The different algorithms can be used as indicated: in pre-training before being applied to a particular experimental system, for real-time optimization and tuning, or a combination of both where the algorithm is pre-trained and subsequently updated during system operation.

FIG. 3. Illustration of machine-learning strategies for optimization and self-tuning of ultrafast fibre lasers using control of intra-cavity elements via a feedback loop and control algorithm. **A**. Training phase where control electronics acting e.g. on the polarization state (EPC: electronic polarization controller) sweep the parameter space to map different operating states of the laser to be used as inputs to the control algorithm (see Fig. 2). Guidelines for algorithm and parameter selection are given in Box~1. In the case of a search algorithm, the training phase is not necessary. Output characteristics are measured by diagnostics tools such as optical spectrum analyser (OSA), fast photodiode (PD) and oscilloscope (OSC), or radio-frequency spectrum analyser (RFSA) and subsequently used as input to the control algorithm. **B**. Machine learning assisted operation where the laser operation is measured in real-time and fed into the control algorithm.

FIG. 4. Machine learning applications in Ultrafast Photonics. **A.** Pulse compression. **Aa**. Optimization procedure. **Ab**. Convergence comparison between neural network and evolutionary algorithm. **Ac**. Compressed pulse FROG. **B**. Controlled nonlinear propagation. **Ba**. Schematic. **Bb** and **Bc**. Examples of customized supercontinuum spectra. **C**. Pulse reconstruction using convolution neural network. **Ca**. Architecture. **Cb**. Reconstructed FROG. **Cc**. Reconstructed pulse. **D**. NLSE solution using a neural network. **Da**. Pulse evolution (top) and comparison of predicted and exact solutions (bottom) at three particular points (dashed lines). **Db**. Kuznetsov-Ma (left) and Akhmediev breather (right) dynamics showing expected evolution (top), predicted evolution (middle), and relative difference (bottom). **E**. Modulation instability. **Ea** Simulated spectra (network input) and **Eb** temporal profiles (network output). **Ec**. Network schematic for correlation of spectral and temporal characteristics. **Ed**. PDF of predicted temporal intensity based on experimental spectra (dashed red line) compared with simulated PDF (blue line). Panel A adapted with permission from REF [42], OSA. Panel B is adapted from REF [47], Springer Nature Ltd. Panel C adapted with permission from REF [57], OSA, Panel Da adapted with permission from REF [67], Elsevier. Panel Db adapted with permission from REF [69], APS. Panel E adapted from REF [75], Springer Nature Ltd.