PHOTONICS Research

Sensing in the presence of strong noise by deep learning of dynamic multimode fiber interference

LINH V. NGUYEN,[1,*] CUONG C. NGUYEN,[2] GUSTAVO CARNEIRO,[2] HEIKE EBENDORFF-HEIDEPRIEM,[1,3] (iD) AND STEPHEN C. WARREN-SMITH[1,3,4] (iD)

[1]*Institute for Photonics and Advanced Sensing and School of Physical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia*
[2]*Australian Institute for Machine Learning, The University of Adelaide, Adelaide, SA 5005, Australia*
[3]*Australian Research Council Centre of Excellence for Nanoscale BioPhotonics, The University of Adelaide, SA 5005, Australia*
[4]*Future Industries Institute, University of South Australia, Mawson Lakes, SA 5095, Australia*
*Corresponding author: linhnguyen.research@gmail.com

A new approach to optical fiber sensing is proposed and demonstrated that allows for specific measurement even in the presence of strong noise from undesired environmental perturbations. A deep neural network model is trained to statistically learn the relation of the complex optical interference output from a multimode optical fiber (MMF) with respect to a measurand of interest while discriminating the noise. This technique negates the need to carefully shield against, or compensate for, undesired perturbations, as is often the case for traditional optical fiber sensors. This is achieved entirely in software without any fiber postprocessing fabrication steps or specific packaging required, such as fiber Bragg gratings or specialized coatings. The technique is highly generalizable, whereby the model can be trained to identify any measurand of interest within any noisy environment provided the measurand affects the optical path length of the MMF's guided modes. We demonstrate the approach using a sapphire crystal optical fiber for temperature sensing under strong noise induced by mechanical vibrations, showing the power of the technique not only to extract sensing information buried in strong noise but to also enable sensing using traditionally challenging exotic materials.   © 2021 Chinese Laser Press

https://doi.org/10.1364/PRJ.415902

1. INTRODUCTION

Changes in the optical path length (OPL) of the guided modes in an optical fiber can be induced by various external perturbations such as temperature, strain, pressure, and biochemical binding [1–5]. This is the underlying principle of many optical fiber sensors (OFSs), particularly intrinsic type OFSs, to date. For the same reason, intrinsic OFSs are cross-sensitive to any other environmental parameters that affect the OPL beyond the intended measurand. This ubiquitous issue for intrinsic OFSs, indeed any sensor, requires careful management or compensation [3–5].

Most intrinsic OFSs developed to date are based on converting the change in an external measurand to a measurable change in OPL of the waves propagating in the fiber. Calibration of these intrinsic OFSs typically involves a single variable, where the change in the measurand is mapped to one measurable optical signal, be it a change in optical intensity or the shift of a resonant wavelength or interference fringe. Even for a spectral-based sensor such as a fiber Bragg grating (FBG), the calibration does not utilize the whole spectrum in the sense that it only focuses on the spectral shift of one

dominant optical feature, the reflected Bragg wavelength. Indeed, much of the effort in developing intrinsic OFSs is to create interference structures such as gratings or cavities so that there are such discrete resonant wavelengths. However, when changes from other perturbations are also encoded in the wavelength shift, the measurement becomes nonspecific, and such ambiguities need to be accounted for in real time. While having been implemented in various forms, the general approach to active compensation against undesired perturbations is to simultaneously measure both the intended measurand and the other competing perturbations using multiple sensor heads. Each sensor head is designed to yield different sensitivity to each measurand so that the calibration maps several optical features to each measurand to discriminate the cross-sensitivity between them [6,7]. A search with the keywords "simultaneous measurement fiber sensor" in Google Scholar returns roughly 400,000 papers for various dual-parameter sensing schemes [8].

The challenge becomes further pronounced when utilizing multimode optical fibers (MMFs) for sensing. Despite being otherwise deterministic optical systems [9], transmission through MMFs is both complex, due to the interference between many propagating optical modes, and dynamic, due

to their high sensitivity to environmental disturbances. That is, any small change in the OPL, such as through temperature, strain, bending, pressure, or biochemical binding, will lead to complex and sensitive changes in the interferometric output. Sensing can be achieved using MMFs, such as by tracking changes in the speckle output of multimode optical fibers [10], but it is often done as an ensemble measurement on the entire specklegram output rather than individual speckle features and is therefore highly susceptible to cross-sensitivity. The complex yet deterministic behavior of the MMF has also been well utilized for imaging using techniques such as optical or digital phase conjugation [11,12], wavefront optimization methods [13–17], or machine learning methods such as convolutional neural networks (CNNs) [9,18]. However, the dynamic behavior of MMFs due to undesired environmental perturbations is a hindrance to both sensing and imaging, as the propagation characteristics will change due to environmental disturbances and thus reduce image quality or cause sensor drift. It is challenging to avoid these impacts due to the complex interactions of many interfering optical modes. While deep neural networks (DNNs) have previously been applied to handle dynamic changes in the MMF, such as for spectrometry [19] and pulse characterization [20], they are fundamentally analyzing the optical input rather than quantifying the dynamic changes that are occurring to the optical fiber itself, that is, sensing.

In this work we seek to exploit both the complex and highly sensitive nature of multimode optical fibers to achieve a new paradigm in optical fiber sensing: specific measurement even in the presence of strong noise from other perturbations. Working in the wavelength domain, using a broadband input to an MMF, we vary the measurand of interest and at the same time apply random strong noise on the fiber to generate a large number of complex output spectra from the MMF. These spectra, encoded with both the measurand information and noise, are used together with the measurand labels to train a DNN to statistically relate the change of the measurand to complex changes in the output spectrum of the MMF while discriminating the noise.

The advantage of this approach is that the DNN maps the entire complex optical spectrum to the measurand change rather than focusing on single or several principal spectral features, such as an FBG peak, eliminating the need for such resonant structure for sensing. MMFs particularly present a unique platform to consider this DNN strategy due to their property of guiding many hundreds or thousands of discrete modes that can each contain important and unique information related to sensing.

We demonstrate this concept for the specific case of temperature sensing using sapphire crystal optical fiber (SOF), a single crystal fiber that has great potential for extreme temperature sensing, but where progress has been hampered by the complexity in sensing using this unstructured and thus highly multimoded waveguide [21–23]. The SOF is therefore particularly suitable for demonstrating the power of our proposed approach, but our technique can be applied, in principle, to any measurand in the presence of any environmental noise using any multimode optical waveguide.

2. PRINCIPLE OF OPERATION AND EXPERIMENTAL DESIGN

The output intensity I from an MMF, assuming coherence is maintained, is a superposition of N guided modes and can be expressed as

$$I(x,y,\lambda) = \left| \sum_{i=1}^{N} a_i \hat{\vec{e}}_i(x,y) e^{i\frac{2\pi}{\lambda} n_i^{\text{eff}} L} \right|^2, \quad (1)$$

where a_i, $\hat{\vec{e}}_i(x,y)$, and n_i^{eff} are the modal amplitude, orthonormal electric field distribution, and effective refractive index of the ith mode, respectively; λ is the free space wavelength, and L is the length of the fiber. Equation (1) represents a speckle pattern at the output of the MMF for a given wavelength. One could measure the output speckle pattern of the MMF using a camera, but this is unsuitable for temperature sensing, the focus of our work, as imaging optics would need to be contained within a high-temperature furnace. We choose to utilize wavelength domain interference, and for this the orthogonality of the modes must be broken through spatial filtering, such as by splicing the MMF to a single-mode fiber (SMF) [24]. Through the coupling of the many modes of the MMF into the SMF, a complex interference spectrum is formed with the number of features increasing as the number of modes N increases. That is, each pair of modes in the MMF creates an interference that is transposed into a spectral interference feature after coupling into the SMF. This is equivalent to the high-dimensional data suitable for training a DNN [25].

The output interference spectrum is highly sensitive to any environmental changes on the optical fiber such as temperature, strain, and pressure, or a combination of these and others, as they affect the OPL of the multimode fiber's modes and thus the phase of the interference spectrum. For example, a temperature change will change the refractive index of the MMF material through the thermo-optic effect and thus change the effective index of the propagating modes. If the environmental changes have different impacts on different modes, albeit an intended measurand or unwanted noise, it is in principle possible to decouple these effects through analysis of the interference spectrum. In the case of temperature, each mode effective index responds differently, as the optical field has varying degrees of overlap with the MMF materials (e.g., sapphire core and air cladding), while mechanical effects such as strain largely impact the MMF length. While these changes to the output spectrum are highly sensitive and complex, in principle they contain rich information on changes that have occurred along the fiber length, and our task is to find an accurate method to extract such information.

Figure 1 shows a schematic diagram of our proposed approach for specific measurement under strong noise using MMF and DNN. The multimode interference in the wavelength domain was realized via a single-mode/multimode/single-mode fiber configuration [Fig. 1(a)]. Broadband light from a lead-in SMF is decomposed into the multiple propagation modes of the MMF. At the end of the MMF a wavelength-dependent speckle pattern is formed, which then couples to the lead-out SMF, creating a complex wavelength-dependent interference spectrum at the output. Due to the large number of modes in the MMF, this interference is highly sensitive to
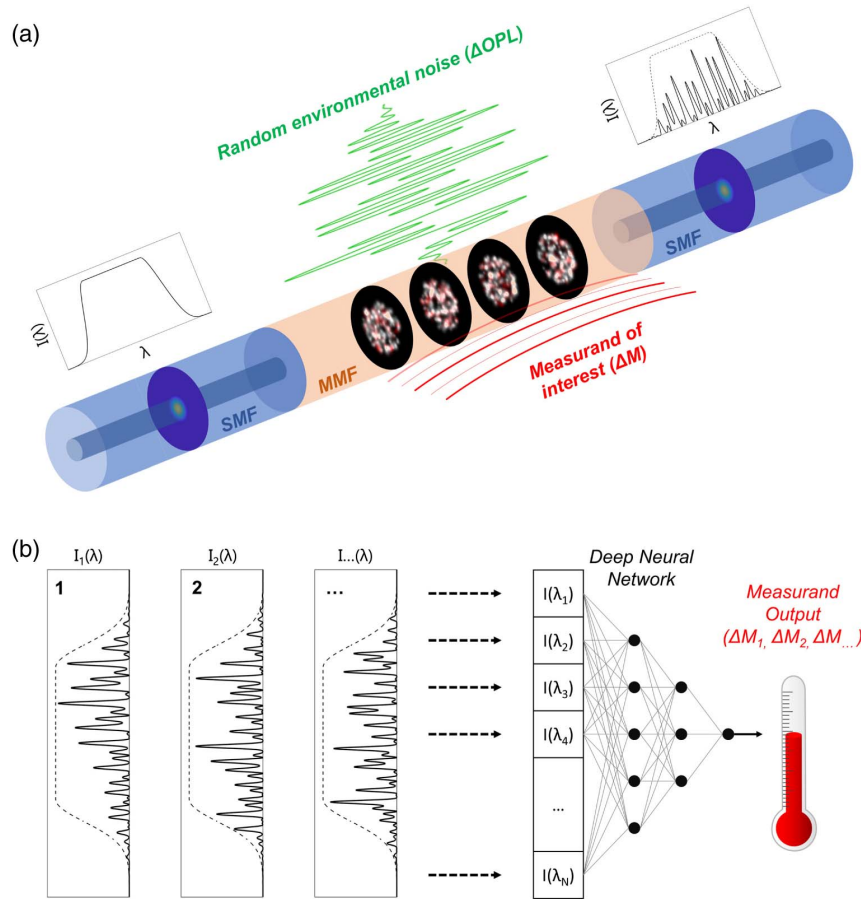
Fig. 1. Schematic diagram of specific measurement under strong noise using MMF and DNN. (a) An SMF-MMF-SMF structure is used to realise complex interference spectra from MMF with a broadband input. The sensing MMF is subjected to both changes in the measurand and strong noise to create complex changes in the output interference spectrum. The speckle images in the MMF conceptually show the impact of the measurand (red), which may be relatively small compared to the noise (white). (b) Various spectra together with their corresponding measurand labels are used to train a DNN. Once trained, the DNN is able to predict the measurand value buried within each interference spectrum.

the splicing condition between the fibers (e.g., tolerance in sensor fabrication) or any perturbations on the MMF. By using the DNN [Fig. 1(b)], even if the spectral change due to a chosen measurand is buried under strong spectral noise created by other environmental perturbations, it is still possible to extract the measurand information unambiguously. On the other hand, this task is extremely challenging if using traditional calibration techniques as shown in Fig. 2.

In our proof-of-concept experiment we used a 200 mm length of multimoded SOF spliced between two standard single-mode fibers (SMF28e). Details of the experimental setup are presented in Appendix A, but briefly, transmission interference spectra were measured using a swept wavelength optical sensor interrogator (National Instruments PXIe-4844) operating at 1 Hz. The spectral resolution of the interrogator was 4 pm over 1510–1590 nm, which was downsampled to 1000 spectral data points (0.8 nm spacing). The SOF was mounted loosely inside a stainless steel tube, which was vibrated with a white noise vibration spectrum using a mechanical shaker. The mounted SOF was placed into a temperature-controlled tube furnace and the temperature was recorded using a thermocouple adjacent to the SOF.

The strong noise induces significant changes in the OPL of the SOF modes as shown in Fig. 2(a). Figure 2(b) shows a similar set of spectra under the same noisy conditions at several different temperatures. The impact of the mechanical noise on the sensor makes it difficult to calibrate the wavelength shift or the intensity variation of an individual peak or trough with respect to a change in temperature. Figures 2(c) and 2(d) show the predicted temperature from the many spectra at different temperatures in the presence of strong mechanical noise using traditional calibration techniques. Figure 2(c) was obtained through direct peak tracking [26], which can be improved by applying a centroid function to determine the peak position [27]. These methods can be considered the most intuitive to apply but are not expected to perform well with complex interference spectra. A more appropriate method to extract shifts in interferometric data is to use the Fourier phase-shift method [28–31] as shown in Fig. 2(d). Here the phase of the dominant Fourier component was tracked, which corresponds to a free spectral range of 3.6 nm for the data in Fig. 1(b). It is evident that these traditional methods suffer poor accuracy in the presence of strong noise, with only 15%, 33%, and 40% of values determined within ±1°C for the peak tracking, centroid, and
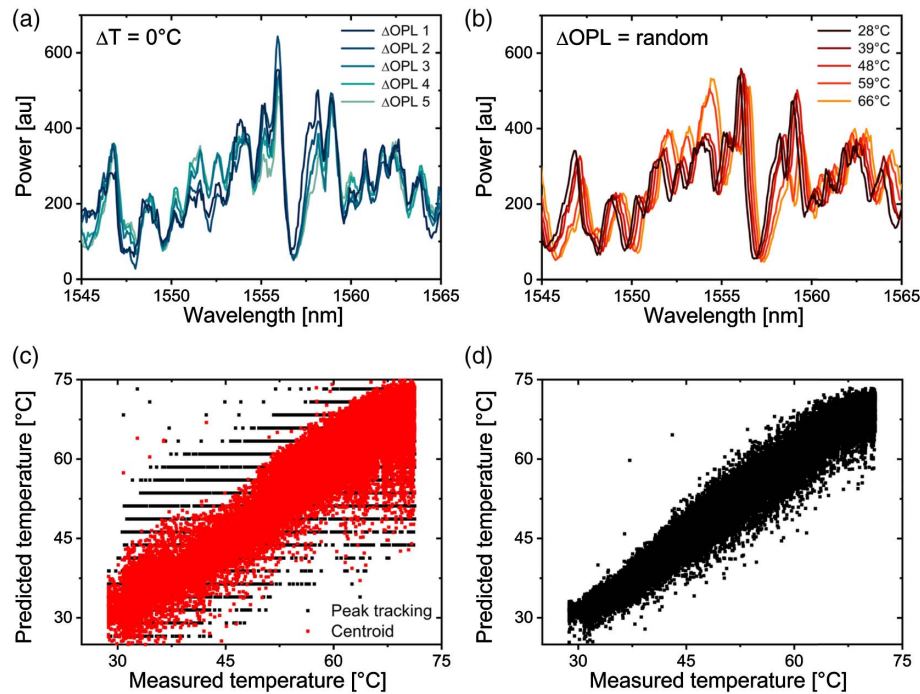
Fig. 2. Complex cross-sensitivity problem in specific measurement with MMF-based sensors under OPL noise. (a) Transmission interference spectra from the SOF at a fixed temperature under white noise in OPL created by shaking the loosely mounted SOF. (b) Interference spectra at several temperatures under OPL noise. (c) Predicted (calibrated) temperature by peak tracking (black) and after applying a centroid function to improve the accuracy (red). (d) Predicted (calibrated) temperature obtained using a Fourier phase-shift technique. The measured temperature in (c) and (d) was obtained using a reference thermocouple.

Table 1. Temperature Measurement Accuracy for Different Calibration Methods[a]

	±1°C	±2°C	±5°C
Peak tracking	15%	30%	70%
Centroid	33%	59%	89%
Fourier phase shift (linear)	33%	60%	96%
Fourier phase shift (quadratic)	40%	70%	97%
Deep learning	99%	100%	100%

[a]For the peak tracking, centroid, and Fourier methods a linear fit has been applied. In addition, a quadratic fit was also applied for the Fourier method to account for nonlinearities in the response seen in Fig. 2(d).

Fourier methods, respectively. On the other hand, as will be discussed in the following sections, a trained DNN can significantly improve the temperature measurement on the same data as shown in Table 1.

3. DEEP LEARNING OF THE DYNAMIC MULTIMODE INTERFERENCE

The results above confirm that without means to shield or compensate OFSs for such cross-sensitivity, calibration for the intended measurand becomes impossible. In addition, when it comes to using the multimode sensor architecture, arguably the simplest optical fiber sensing structure, any deviation in sensor preparation such as splicing and cleaving will lead to complex changes in the interference spectrum due to the involvement of many modes. These deviations are noise in their

own right. Therefore, a calibration technique that works directly with noisy data is of great advantage, as all the complexity in materials and optical engineering to mitigate cross-sensitivity is shifted to software, a much more scalable and adaptable approach.

Supervised deep learning or DNN can be thought as a highly complex fitting function that maps a high-dimensional raw input, such as an image, to an output, such as another image, the class of an object, or a measurement value [25]. The training is typically performed by showing the DNN many example inputs and requesting it to optimize a loss function, that is, the difference between ground truth and the output of the DNN model. Once the loss function is minimized for a large number of example inputs the model is considered trained and should be able to generalize to inputs it never saw during training. While DNNs have been investigated for decades [32], including early applications of DNN in MMF [33–35], recent advances in computing power, availability of data, and advanced network architectures have spawned the use of DNN for a very wide range of successful applications [25]. This includes recent promising results in reconstructing or recognizing image transmission through scattering media such as MMF [9,18,36]. In this case, the DNN is trained to understand the relation between the MMF output and input (mapping the output speckle to the original input image) where the MMF is protected from environmental perturbations. The learning capability, given a fixed network structure and data set, deteriorates with respect to the MMF length due to perturbations on the MMF such as temperature and vibration,

leading to an unstable speckle pattern at the distal end [18]. Therefore, such learning can be considered static in the sense that the transmission medium is fixed, and the training data is generated by many different input-speckle pairs.

Rather than training the DNN with many different labelled images, here we use a fixed optical input to the MMF and rely on the change of a specific measurand together with random noise along the MMF to generate a large number of different outputs, which are then used together with the measurand label to train the DNN. In this case, without prior knowledge of the light propagating in the MMF, the DNN is trained to statistically understand the relationship between the measurand and the distorted interferometric output from the MMF while ignoring the noise. In other words, the DNN is taught to learn the dynamic multimode interference with respect to an intended perturbation buried under the noise.

Figure 3 shows a schematic diagram of the DNN architecture we used for learning the relation between the transmission spectra and temperature. Here the input is MMF interference spectra inclusive of both temperature changes and noise as depicted in Fig. 1, and the DNN is trained to understand such changes buried in the OPL noise induced by strong mechanical vibration. This is analogous to the way in which DNN networks are trained to understand a cat or a dog in various lighting and background settings (noise). We used a multilayer perceptron (MLP) [37] consisting of four hidden layers, each having 512, 256, 128, and 64 nodes, respectively, resulting in 684,096 trainable parameters. The output of each layer was activated by the rectified linear unit (ReLU) [38]. No regularization, dropout, or batch normalizations were used. The loss function optimized in this regression case was the standard mean square error (MSE). An Adam optimizer [39] with a fixed learning rate of 10^{-3} was used to minimize the loss. All the code was conducted using Keras—an open-source neural network written in Python [40].

While there are numerous open source deep learning architectures available, they are generally convolutional neural networks (CNNs) designed to work with 2D images that typically have a high level of spatial correlation between neighboring pixels. These large architectures are often trained using the large data set ImageNet [41], which consists of 14 million daily life images. In contrast, the spectral data in our work is a one-dimensional array containing intensity values at different wavelengths. The wavelength spacing in our experiment was set at 0.8 nm, which we can compare to an estimate of the correlation bandwidth for a multimode optical fiber given by $\delta\lambda \approx \lambda^2/(2\Delta nL)$, where Δn is the difference between the highest and lowest effective indices supported by the optical fiber [42]. For a 20 cm length of SOF with a refractive index of 1.7 surrounded by air ($n = 1.0$) the correlation bandwidth is estimated to be 8.6 pm, much less than our wavelength spacing. This means that each intensity value at one wavelength (superposition of many modes at said wavelength) is decorrelated from its neighbors and can be treated as an independent feature. Therefore, an MLP that is designed based on matrix multiplication to describe the interaction between all features within a layer is the preferred architecture for the spectral data obtained from highly multimode interference in our work. Further comparison with results obtained from training and testing the same data sets with a CNN to test this hypothesis is presented in Section 4.B.

To train and evaluate the MLP, the data set was split into three subsets: training, validation, and testing. The training and validation subsets were used to train and cross-validate to fine-tune hyperparameters such as learning rate and mini-batch size, while the testing subset was reserved for evaluation purposes.

The MLP was trained following a standard iteration process.

(i) Mini-batches of spectra were randomly sampled from the training subset and passed forward through the MLP, producing predicted temperature values.

(ii) The MSE between the predicted and the actual temperatures was calculated.

(iii) The error was minimized by adjusting the network's weights following the standard backpropagation.

(iv) When all mini-batches in the training data set were completed, corresponding to 1 epoch, the training set was re-randomized and the next iteration of the training commenced.

(v) This process was repeated until the error or loss converged to a minimum value.

The training was carried out using approximately 50,000 SOF transmission spectra inclusive of random noise and their
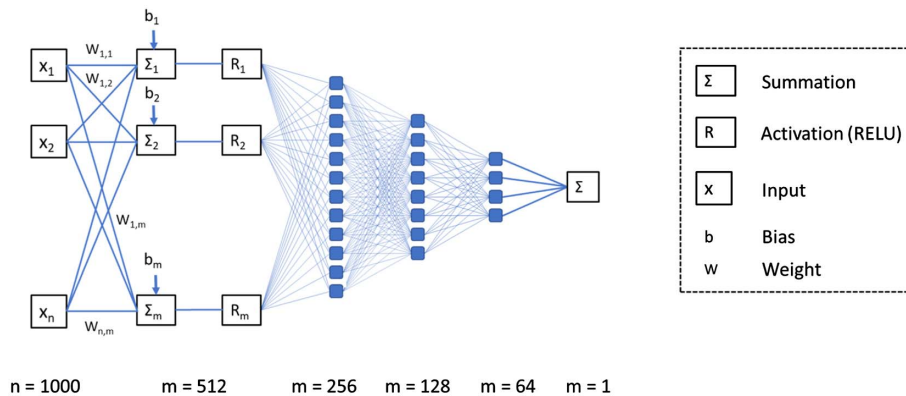

Fig. 3. Schematic diagram of the multilayer perceptron (MLP) architecture used in this work. After being trained using thousands of different noisy MMF interference spectra with their corresponding temperature labels, the trained model takes an MMF spectrum as its input and produces a predicted temperature value at its output.

corresponding temperature labels. Training took approximately 7 h on an Intel Core i7-7700HQ CPU at 2.8 GHz and 32 Gb RAM until the loss converged. During the training, the prediction accuracies on the training subset and the held-out validation subset were monitored after every epoch to fine-tune hyperparameters, if necessary, to prevent overfitting.

To evaluate the trained MLP, the spectra in the testing subset were passed through the network to predict the corresponding temperatures. The obtained temperatures were then compared with the ground-truth temperatures to assess the accuracy of the predictions made by the trained MLP.

4. RESULTS AND DISCUSSION

A. Temperature Sensing under Strong Noise
Figures 4(a) and 4(b) show the reference (thermocouple) measured temperatures and their corresponding predicted values from the trained model. Two different data sets were used for training as shown in Table 2. The modal accuracy is defined as the fraction of model-predicted values that are within ±1°C or ±2°C of the measured values. The trained MLP successfully maps almost all the values in the interference spectrum to their corresponding temperature while discriminating the effect from the OPL noise. The better accuracy for the T1 data set, reaching a ±1°C accuracy of 99%, is attributed to the finer temperature variation in the training data compared with that of the T2 data set.

Figures 4(c) and 4(d) show the inter-testing results where the DNN is trained with one data set and tested on the other.

Table 2. Data Set Used for Training and Testing

Data Set Label	Temperature Range (°C)	Ramp Rate (°C/h)	Total Spectra	Data Collection Time (h)
T1	29–75	1	84,000	23.3
T2	29–630	15	78,450	21.8

For the case of training with T1 and testing with T2, shown in Fig. 4(c), the accuracy is very poor, since there are many temperature values that are out of the training range. When the testing data T1 is within the training data T2, as shown in Fig. 4(d), the accuracy is better but still significantly less than the case of testing with the same data set T2 [e.g. ±2°C accuracy of 92% in Fig. 4(d) compared with 100% in Fig. 4(a)]. This is still significantly better than the Fourier shift method result (70%). We also note that the amount of training data in the temperature range of interest (29°C–75°C) is significantly less for Fig. 4(d) compared to Fig. 4(a) at approximately 6000 spectra compared to 78,450. This is because the 29°C–75°C region is only a small subset of the total T2 data set. One would therefore expect that the accuracy could be improved by increasing the number of spectra in the T2 training data set.

In this particular case of temperature sensing using SOF, since the furnace temperature includes OPL change along the entire SOF while the temperature labels are collected at a single point, the T1 and T2 data sets are likely shifted in their thermal distributions with respect to temperature values, and thus they would not work well in an interplay manner.
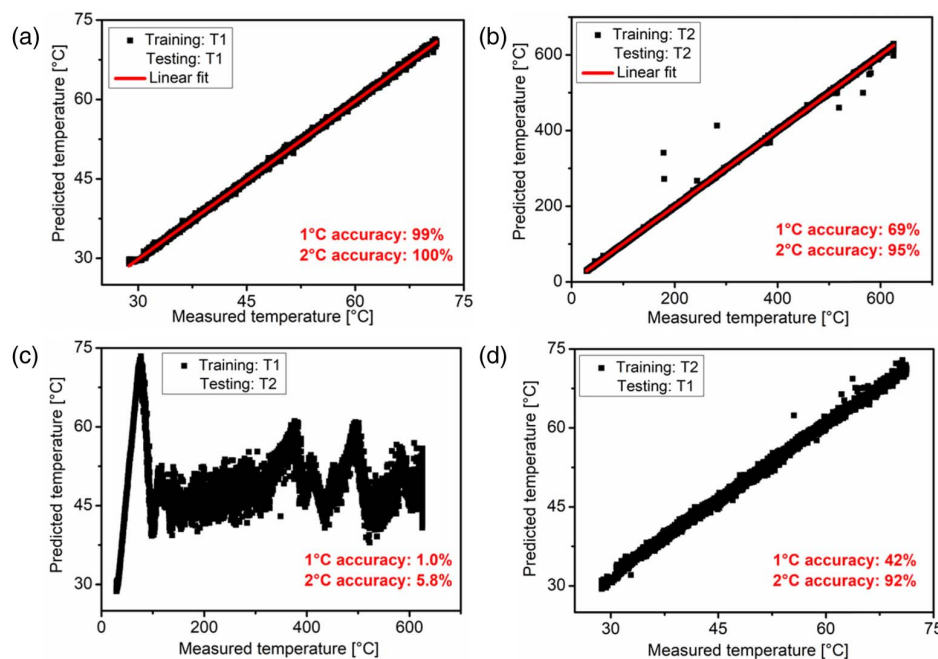

Fig. 4. Comparison of temperatures predicted from the trained MLP using SOF interference spectra as input and measured temperatures under OPL noise. (a) Training and testing on the same temperature range using the short temperature range from 30°C to 70°C (T1) with 8400 testing examples (data points). (b) Training and testing on the same temperature range using extended temperature range from 30°C to 630°C (T2) with 7845 testing examples. (c) Same procedure but training and testing on different data sets, the MLP is trained on the short temperature range and tested on the extended temperature range, where many temperature values are out of the training range. (d) The MLP is trained on the extended temperature range and tested on the short temperature range. In this case all the testing temperatures are predicted within the range of the training set but with less accuracy compared with (b) due to a shift in the data set.
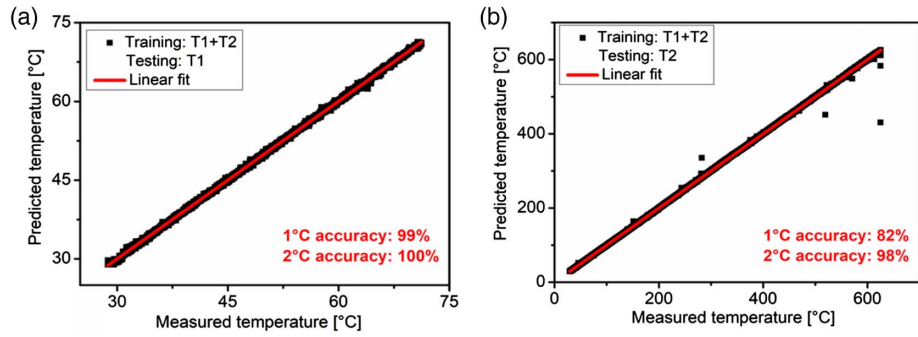
Fig. 5. Mitigating the domain shift issue using a combined data set T1 + T2. (a) The testing set was T1 using 8400 example spectra. (b) The testing set was T2 using 7845 example spectra. The test results were improved for both cases.

These results highlight a well-known issue of machine learning generalization where the testing data belongs to a different data distribution compared with the training data [43–46].

While mitigating the generalization issue with advanced deep learning techniques is not within the scope of this paper, readers are referred to reviews of this active area of machine learning research [47–49]. Nevertheless, the most straightforward approach is to collect more diverse data. In the case of optical fiber sensing, it implies that the noise needs to include diversity in OPL variations that mimic the changes due to many other perturbations, as well as obtaining measurand labels in a localized manner (i.e., distributed measurement) rather than a single point. Figures 5(a) and 5(b) show the improved results where the training was done on a more diverse data set, which is the combination of T1 and T2, and tested with constituent data sets. However, it is often very time consuming and expensive to have a large and diverse data set readily collected, and hence a more practical approach is to simply deploy a model trained on available data and continuously fine-tune it with newly collected data [50].

B. Temperature Sensing: Comparison between MLP and CNN Architectures

A 1D convolutional neural network [10] with a similar number of parameters (see Appendix B) was also trained and tested using the same data sets to determine whether the model can be improved by taking into account the wavelength locality. As discussed previously in Section 3, the SOF supports many different guided modes, yielding a decorrelated interference spectrum that does not have a regular form. Therefore, it is not expected that the CNN can perform with greater accuracy compared to the MLP, in contrast to traditional image recognition problems. Figure 6 compares the 1°C accuracy between the MLP and CNN models for the cases of T1, T2, and the combined data set T1 + T2. The MLP is found to consistently perform better than its CNN counterpart for a similar number of network parameters.

The MLP and CNN architectures are two different kinds of neural networks used to model different types of data [51]. MLPs are designed based on matrix multiplication to describe the interaction between all features within a layer, allowing the network to work with any type of data. The trade-off is the linearly increasing number of parameters used. In contrast,
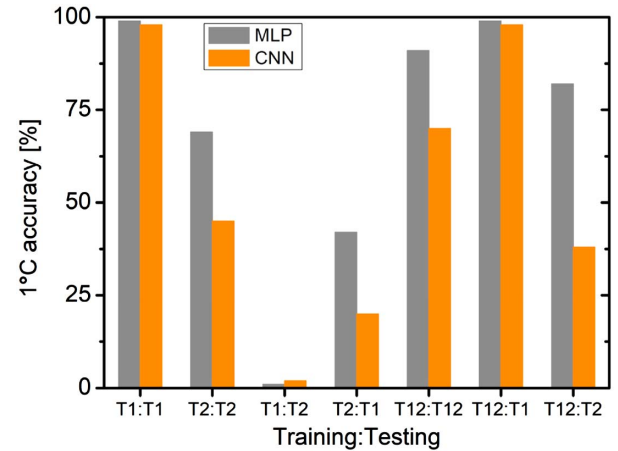

Fig. 6. Comparison between MLP and CNN architectures with a similar number of parameters. Training and testing were done in the same manner as in Figs. 4 and 5. The CNN produces inferior predictions compared with its MLP counterpart, indicating that the raw interference spectra from the highly multimode SOF do not possess significant wavelength locality.

CNNs rely on the convolution operation that models sparse interaction between features, resulting in fewer parameters. Due to the spatial nature of the convolution operation, CNNs are likely to perform better on data where the correlation is strong between adjacent features and gradually reduces for features that are further away. In our case, where the interference spectrum arises from the highly multimode SOF with many modes involved, there should be weak localized structure in the wavelength spectrum, and the superposition of all modes at a certain wavelength can be treated as an independent feature to its neighbor. Therefore, the MLP is preferable over the CNN in learning this type of data, due to both the increased accuracy shown in Fig. 6 and the relative simplicity of the network.

5. CONCLUSIONS AND OUTLOOK

We have proposed the use of deep learning to learn the dynamic behavior of multimode fiber interference in order to achieve specific measurement even under strong noise, without shielding the sensor from, or compensating for, cross-sensitivity

from undesired perturbations. We have achieved a proof-of-concept demonstration for the specific example of mapping the interference spectrum of an SOF and temperature, under induced OPL noise from mechanical shaking. From here our proposed technique can be expanded to any type of measurand and OPL noise provided a sufficiently diverse data set is obtained. The technique can also be readily extended to other MMF optical waveguides, again, provided a training data set is obtained for that device.

In this proof-of-concept demonstration we have focused on a single optical fiber sensor, whereby the training data and test data have been measured with the same device. A key future investigation will be to understand the generalization of our model for other nominally identical devices within prescribed tolerances. One would expect that model generalization could be achieved in a brute force approach by collecting diverse training data for many similar, but slightly different, optical fiber sensors. Understanding and solving this problem in a practical way would be key to the successful implementation of our method in an industrial setting.

This work opens an entirely new direction in optical fiber sensing in which complexity in hardware to combat cross-sensitivity or noise is shifted into software using intelligent sensor calibration that exploits advances in machine learning.

APPENDIX A: EXPERIMENTAL SETUP

Figure 7 shows a schematic diagram of the experimental setup that was used to collect SOF interferometric data inclusive of OPL noise and temperature effect for training and testing the DNN architectures. A 200 mm length of SOF was spliced with lead-in and lead-out SMFs and inserted into a stainless steel (SS) tube of 7 mm inner diameter with both ends glued such that the SOF was loosely hung inside the SS tube. The SOF was purchased from Micromaterials and has a quoted diameter of 70 µm.

Given the diameter and refractive index of the SOF, the fiber is estimated to support approximately 100,000 modes. However, it should be noted that the SOF is made via a crystal growth process that produces a diameter variation along the fiber, and thus the number of supported mode numbers is expected to vary along the fiber as well. The SS tube was connected to a mechanical shaker (Bruel & Kjear LDS V455) that was driven with white noise vibration from 20 to 2000 Hz with a root-mean-squared acceleration of $G_{\mathrm{rms}} = 3.0g$.

Table 3.   Example Data Set

TC (°C)	I_1 (AU)	I_2 (AU)	I_3 (AU)	...	I_{1000} (AU)
28.722	346.607	386.807	430.726	...	285.396
37.895	111.564	124.290	151.859	...	234.433
33.059	295.809	363.931	385.505	...	286.972
31.913	262.195	301.755	343.645	...	291.099
33.059	295.809	363.931	385.505	...	286.972
38.173	295.809	207.829	263.489	...	271.466
49.200	172.730	92.328	103.878	...	191.467
64.408	81.229	223.914	215.993	...	125.078
70.939	242.774	223.597	162.985	...	114.785
41.516	260.047	97.779	98.382	...	231.725

This created random movement of the SOF within the SS tube while the temperature was independently varied.

The transmission spectrum of the SMF-SOF-SMF structure was recorded using an optical sensor interrogator (National Instrument PXle-4844). The interrogator has a wavelength resolution of 0.004 nm, which was swept at 1 Hz over a wavelength bandwidth of 80 nm so that a full range spectrum contains 20,000 intensity values. No polarization optics were used as any polarization state output from the interrogator would in any case be scrambled in the highly multimoded vibrating sapphire optical fiber. To reduce the number of dimensions in the input data, the wavelength resolution was reduced 20× to 0.08 nm, and thus the resultant spectra used for training and testing the DNN were composed of 1000 intensity values. A thermocouple was placed at the center of the SOF, and the temperature value was recorded in synchronization with the spectrum acquisition. Each data point is therefore represented as a 1001-dimensional vector whose first element is the temperature label (output) and the remaining 1000 elements are the spectrum (input). A random selection of 10 data points (spectra together with their temperature labels) is shown in Table 3. The intensity values in each spectrum were normalized to values between 0 and 1 for training.

APPENDIX B: 1D CONVOLUTIONAL NEURAL NETWORK

A schematic diagram of the one-dimensional convolutional neural network (1D CNN) used to train the same data sets that led to the comparison with the MLP in Fig. 6 is shown in Fig. 8. It has hidden layers consisting of six convolutional
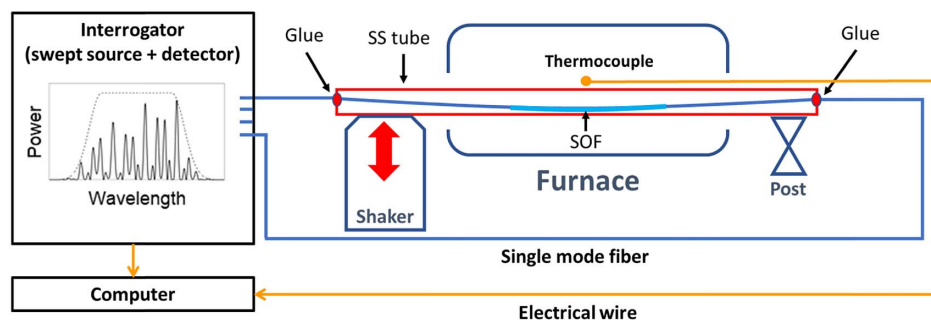
Fig. 7.   Experimental setup. The SOF was mounted loosely inside the SS tube so that it vibrated within the SS tube. The shaker was driven by a white noise vibration profile. The SS tube was placed centrally within a tube furnace and the temperature referenced with a thermocouple.
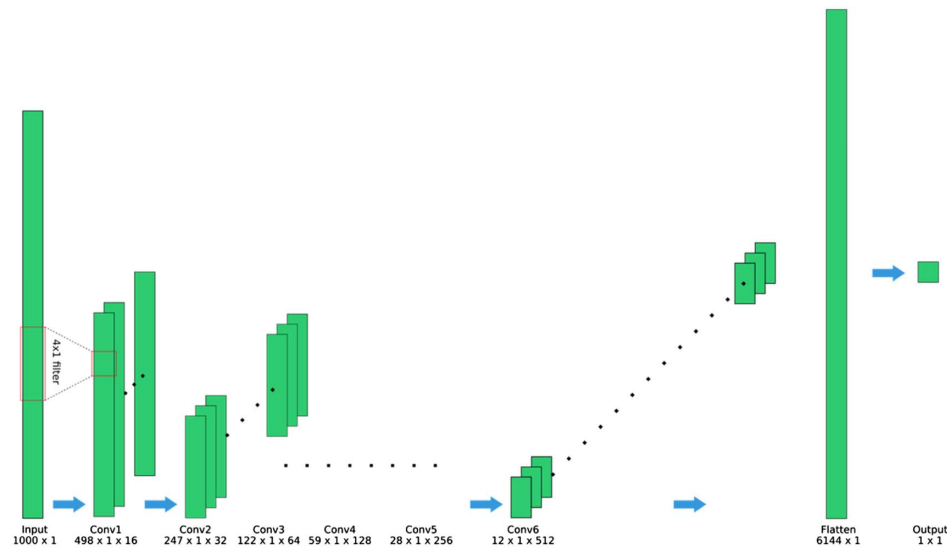
Fig. 8. Schematic diagram of the 1D CNN architecture.

modules and one fully connected layer with a total of 707,601 trainable parameters, comparable to that of the MLP. Each module uses a 4×1 convolutional kernel, followed by batch normalization, ReLU, and 2×1 max pooling. The number of channels in each layer is doubled as their sizes are halved through down-sampling. The output of the final convolutional module was flattened onto a single vector of 6144 elements and passed through the last fully connected layer to output predicted temperature. The training of the 1D CNN was similar to that of the MLP.

Disclosures. S.C.W.S. is a director of HT Sensing Pty. Ltd., a company that manufactures optical fiber sensors. HT Sensing Pty. Ltd. did not contribute to or participate in this research in any way.

REFERENCES AND NOTES

1. D. A. Krohn, T. W. MacDougall, and A. Mendez, *Fiber Optic Sensors: Fundamentals and Applications*, 4th ed. (SPIE, 2015).
2. B. Lee, "Review of the present status of optical fiber sensors," Opt. Fibre Technol. **9**, 57–79 (2003).
3. X. Wang and O. S. Wolfbeis, "Fiber-optic chemical sensors and biosensors (2015–2019)," Anal. Chem. **92**, 397–430 (2020).
4. X. He, C. Ma, S. Wang, Z. Wang, and L. Yuan, "Pressure vector sensor based on an orthogonal optical path Sagnac interferometer," Opt. Express **28**, 7969–7979 (2020).
5. Q. Tian, G. Xin, K.-S. Lim, Y. He, J. Liu, H. Ahmad, X. Liu, and H. Yang, "Cascaded Fabry-Perot interferometer-regenerated fiber Bragg grating structure for temperature-strain measurement under extreme temperature conditions," Opt. Express **28**, 30478–30488 (2020).
6. V. Bhatia, D. Campbell, R. O. Claus, and A. M. Vengsarkar, "Simultaneous strain and temperature measurement with long-period gratings," Opt. Lett. **22**, 648–650 (1997).
7. X. Li, L. V. Nguyen, H. Ebendorff-Heidepriem, D. Pham, and S. C. Warren-Smith, "Simultaneous measurement of temperature and refractive index using an exposed core micro structured optical fiber," IEEE J. Sel. Quantum Electron. **26**, 5600107 (2020).
8. Google Scholar search using "simultaneous measurement fiber sensor," accessed October 25, 2020.
9. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," Light Sci. Appl. **7**, 69 (2018).
10. T. D. Cabral, E. Fujiwara, S. C. Warren-Smith, H. Ebendorff-Heidepriem, and C. M. B. Cordeiro, "Multimode exposed core fiber specklegram sensor," Opt. Lett. **45**, 3212–3215 (2020).
11. A. Yariv, "Phase conjugate optics and real-time holography," IEEE J. Quantum Electron. **14**, 650–660 (1978).
12. I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, "Focusing and scanning light through a multimode optical fiber using digital phase conjugation," Opt. Express **20**, 10583–10590 (2012).
13. T. Čižmár and K. Dholakia, "Shaping the light transmission through a multimode optical fibre: complex transformation analysis and applications in biophotonics," Opt. Express **19**, 18871–18884 (2011).
14. S. Bianchi and R. Di Leonardo, "A multi-mode fiber probe for holographic micromanipulation and microscopy," Lab Chip **12**, 635–639 (2012).
15. Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, "Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber," Phys. Rev. Lett. **109**, 203901 (2012).
16. A. M. Caravaca-Aguirre, E. Niv, D. B. Conkey, and R. Piestun, "Real-time resilient focusing through a bending multimode fiber," Opt. Express **21**, 12881–12887 (2013).
17. M. N'Gom, T. B. Norris, E. Michielssen, and R. R. Nadakuditi, "Mode control in a multimode fiber through acquiring its transmission matrix from a reference-less optical system," Opt. Lett. **43**, 419–422 (2018).
18. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," Optica **5**, 960–966 (2018).

19. R. K. Gupta, R. D. Bruce, S. J. Powis, and K. Dholakia, "Deep learning enabled laser speckle wavemeter with a high dynamic range," Laser Photon. Rev. **14**, 2000120 (2020).

20. W. Xiong, B. Redding, S. Gertler, Y. Bromberg, H. D. Tagare, and H. Cao, "Deep learning of ultrafast pulses with a multimode fiber," APL Photon. **5**, 096106 (2020).

21. J. Wang, B. Dong, E. Lally, J. Gong, M. Han, and A. Wang, "Multiplexed high temperature sensing with sapphire fiber air gap-based extrinsic Fabry–Perot interferometers," Opt. Lett. **35**, 619–621 (2010).

22. T. Habisreuther, T. Elsmann, Z. Pan, A. Graf, R. Willsch, and M. A. Schmidt, "Sapphire fiber Bragg gratings for high temperature and dynamic temperature diagnostics," Appl. Therm. Eng. **91**, 860–865 (2015).

23. D. Grobnic, S. J. Mihailov, C. W. Smelser, and H. Ding, "Sapphire fiber Bragg grating sensor made using femtosecond laser radiation for ultrahigh temperature applications," IEEE Photon. Technol. Lett. **16**, 2505–2507 (2004).

24. X. Zhu, A. Schülzgen, H. Li, L. Li, L. Han, J. V. Moloney, and N. Peyghambarian, "Detailed investigation of self-imaging in large core multimode optical fibers for application in fiber lasers and amplifiers," Opt. Express **16**, 16632–16645 (2008).

25. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).

26. T. Wieduwilt, J. Dellith, F. Talkenberg, H. Bartelt, and M. A. Schmidt, "Reflectivity enhanced refractive index sensor based on a fiber-integrated Fabry-Perot microresonator," Opt. Express **22**, 25333–25346 (2014).

27. S. C. Warren-Smith, R. Kostecki, L. V. Nguyen, and T. M. Monro, "Fabrication, splicing, Bragg grating writing, and polyelectrolyte functionalization of exposed-core microstructured optical fibers," Opt. Express **22**, 29493–29504 (2014).

28. L. V. Nguyen, K. Hill, S. C. Warren-Smith, and T. M. Monro, "Interferometric-type optical biosensor based on exposed-core microstructured optical fiber," Sens. Actuators B **221**, 320–327 (2015).

29. S. Pevec and D. Donlagic, "High resolution, all-fiber, micro-machined sensor for simultaneous measurement of refractive index and temperature," Opt. Express **22**, 16241–16253 (2014).

30. J. W. Silverstone, S. McFarlane, C. P. K. Manchee, and A. Meldrum, "Ultimate resolution for refractometric sensing with whispering gallery mode microcavities," Opt. Express **20**, 8284–8295 (2012).

31. Y. Jiang, "Fourier transform white-light interferometry for the measurement of fiber-optic extrinsic Fabry-Pérot interferometric sensors," IEEE Photon. Tech. Lett. **20**, 75–77 (2008).

32. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbart, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," Neural Comput. **1**, 541–551 (1989).

33. S. Aisawa, K. Noguchi, and T. Matsumoto, "Remote image classification through multimode optical fiber using a neural network," Opt. Lett. **16**, 645–647 (1991).

34. T. Matsumoto, M. Koga, K. Noguchi, and S. Aizawa, "Proposal for neural-network applications to fiber-optic transmission," in Proceedings of 1990 IJCNN International Joint Conference on Neural Networks (1990), pp. 75–80.

35. R. K. Marusarz and M. R. Sayeh, "Neural network-based multimode fiber-optic information transmission," Appl. Opt. **40**, 219–227 (2001).

36. P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, "Transmission of natural scene images through a multimode fibre," Nat. Commun. **10**, 2029 (2019).

37. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction (Springer, 2009).

38. R. H. R. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, and H. S. Seung, "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit," Nature **405**, 947–951 (2000).

39. D. K. Diederik and J. L. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2014).

40. https://keras.io/.

41. http://www.image-net.org/.

42. B. Redding, S. M. Popoff, and H. Cao, "All-fiber spectrometer based on speckle pattern reconstruction," Opt. Express **21**, 6584–6600 (2013).

43. J. Quionero-Candela, M. Sugiyama, A. Schwaighofer, and N. Lawrence, Dataset Shift in Machine Learning (MIT, 2009).

44. R. Alaiz-Rodríguez and N. Japkowicz, "Assessing the impact of changing environments on classifier performance," in Advances in Artificial Intelligence (Springer, 2008), pp. 13–24.

45. M. Wang and W. Deng, "Deep visual domain adaptation: a survey," Neurocomputing **312**, 135–153 (2018).

46. D. Tasche, "Fisher consistency for prior probability shift," J. Mach. Learn. Res. **18**, 1–32 (2017).

47. S. Geman, E. Bienenstock, and R. Doursat, "Neural networks and the bias/variance dilemma," Neural Comput. **4**, 1–58 (1992).

48. O. Bousquet and A. Elisseeff, "Stability and generalization," J. Mach. Learn. Res. **2**, 499–526 (2002).

49. C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," in International Conference on Learning Representation (2017), pp. 1–15.

50. A. Ng, "Machine learning yearning: technical strategy for AI engineers in the era of deep learning," https://www.deeplearning.ai/machine-learning-yearning/.

51. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning (MIT, 2016).