Research Article | Vol. 27, No. 5 | 4 Mar 2019 | OPTICS EXPRESS 7405
Optics EXPRESS

Real-time dynamic strain sensing in optical fibers using artificial neural networks

SASCHA LIEHR,[1,*] LENA ANN JÄGER,[2] CHRISTOS KARAPANAGIOTIS,[1] SVEN MÜNZENBERGER,[1] AND STEFAN KOWARIK[1]

[1]*Bundesanstalt für Materialforschung und -prüfung (BAM), Unter den Eichen 87, 12205 Berlin, Germany*
[2]*University of Potsdam, Department of Computer Science, August-Bebel-Str. 89, 14482 Potsdam, Germany*
**sascha.liehr@bam.de*

Abstract: We propose to use artificial neural networks (ANNs) for raw measurement data interpolation and signal shift computation and to demonstrate advantages for wavelength-scanning coherent optical time domain reflectometry (WS-COTDR) and dynamic strain distribution measurement along optical fibers. The ANNs are trained with synthetic data to predict signal shifts from wavelength scans. Domain adaptation to measurement data is achieved, and standard correlation algorithms are outperformed. First and foremost, the ANN reduces the data analysis time by more than two orders of magnitude, making it possible for the first time to predict strain in real-time applications using the WS-COTDR approach. Further, strain noise and linearity of the sensor response are improved, resulting in more accurate measurements. ANNs also perform better for low signal-to-noise measurement data, for a reduced length of correlation input (i.e., extended distance range), and for coarser sampling settings (i.e., extended strain scanning range). The general applicability is demonstrated for distributed measurement of ground movement along a dark fiber in a telecom cable. The presented ANN-based techniques can be employed to improve the performance of a wide range of correlation or interpolation problems in fiber sensing data analysis and beyond.

1. Introduction

Distributed acoustic sensing (DAS), or distributed vibration sensing (DVS), along optical fibers is being used for a wide range of applications, such as oil and gas pipeline monitoring, perimeter monitoring, rail track monitoring, structural health monitoring and seismic applications [1]. Most DAS principles involve either direct detection or coherent detection of Rayleigh backscattering from coherent optical pulses that are sent into standard telecommunication optical fibers. More recently, optical frequency domain reflectometry has also been enhanced for dynamic strain sensing [2–4]. COTDR-based DAS techniques allow for dynamic high-resolution strain measurement in the nε (nm/m) range over tens of kilometers with a spatial resolution in the meter-range. The technological advancement has been very dynamic, and the scope of DAS applications has widened significantly in recent years. A variety of detection principles have been proposed [5,6]. The simplest approach is based on backscattering intensity change evaluation, does not allow to accurately measure strain changes, and is subjected to signal fading [7–10]. The most common DAS implementations, however, rely on phase demodulation or phase change tracking, which allows for quantitative strain change measurement. A wide range of different phase demodulation principles has been proposed [11–17].

Another group of evaluation techniques is based on correlation analysis to quantify strain changes from the shift of backscattered power signatures. These power signature shift-based approaches have been demonstrated for correlations along three dimensions, or sampling

#355364 https://doi.org/10.1364/OE.27.007405
Journal © 2019 Received 7 Jan 2019; revised 14 Feb 2019; accepted 15 Feb 2019; published 27 Feb 2019

Optics EXPRESS

parameters: (*i*) along the fiber distance axis using chirped optical pulses [18], (*ii*) along the measurement time axis using alternating pulses of different wavelengths [19], and (*iii*) along the pulse wavelength axis by periodically scanning the fiber with pulses within a range of wavelengths [20–22].

The research presented in this paper is based on the latter, the *wavelength-scanning* approach. The wavelength-dependent backscatter signature is thereby shifted by strain or temperature impact. Each new wavelength scan can therefore be compared to an unaltered reference scan and relative strain changes can be calculated from signal shifts for each position along the fiber. The main advantage of this reference measurement-based approach is that interrupted measurements can be resumed without loss of reference, whereas the phase information is generally lost for conventional phase-tracking DAS principles once they are interrupted. Quasi-static strain changes occurring over extended time periods can therefore be measured with high accuracy.

This wavelength-scanning approach has, until recently, only been demonstrated for static measurement [21,22]. To overcome this limitation, we proposed a simple and robust WS-COTDR technique, which conducts fast pulse-to-pulse wavelength scans of the fiber's Rayleigh backscatter signature while directly modulating the laser current [20]. Dynamic measurement with nε strain resolution at kHz repetition rate over several kilometers is achieved for the first time [20]. However, due to computational limitations the correlation-based raw data processing that is necessary to calculate the strain results so far cannot be conducted in real-time. This principle has therefore been restricted to a limited range of applications and later strain calculation from recorded raw measurement data. In this work, we demonstrate that artificial neural networks (ANNs) can substantially decrease the strain computation time by more than two orders of magnitude. We show that using ANNs for strain prediction not only enables the wavelength-scanning principle to be used for real-time applications, but also considerably improves the sensor performance and extends the measurement range.

The applications of ANNs in science and technology are too numerous to list and we refer to review papers and references therein [23,24]. In optical fiber sensing, ANNs have mostly been used for high-level analyses of measurement data, for example for pattern recognition and event categorization in DAS applications [25–27]. A review including other machine learning methods is presented in [28]. Raw data processing with ANNs has so far been demonstrated to improve the prediction of Brillouin frequency spectra shifts [29], discriminate between strain and temperature effects [30], and for classification of principle component analysis results of Brillouin measurements [31]. ANNs have also been used for noise reduction of 2D measurement data for Brillouin [32]. Quantitative predictions by ANNs from arbitrary and more complex patterns have so far not been demonstrated for fiber optic sensing. We demonstrate that Rayleigh signature shifts can be efficiently predicted by ANNs that were trained with synthetic training data.

This paper is structured as follows: The Rayleigh backscatter interference model that is used for synthesized ANN training data generation (Section 2.1) is introduced in Section 2.2. The previously used reference approach, a least mean square (LMS) correlation technique [20], is presented in Section 2.3 and the experimental setup is described in Section 3. An additional ANN subnet that conducts a necessary laser sweep linearization is introduced in Section 4.1. General sensor performance parameters are defined in Section 4.2.1 and the ANN training and hyperparameter selection are summarized in Section 4.2.2. Performance improvement is demonstrated and summarized for: low signal-to-noise relation measurement data (Section 4.3.1), coarser frequency steps (Section 4.3.2), increased distance range (Section 4.3.3), and in terms of reduced computation time (Section 4.3.4). An application example of ground movement along in a telecommunication cable in an urban environment is demonstrated in Section 4.4.

Optics EXPRESS

2. Theory

2.1. Rayleigh backscatter modelling and wavelength-scanning approach

Figure 1 shows a simplified schematic of the method. The WS-COTDR principle is based on Rayleigh scattering from inhomogeneities in the optical fiber which act as a distributed interferometer (see Fig. 1(a)), leading to partially constructive or destructive interference of the backscattered light. Strain effects can, for example, increase the path length between scatterers at a specific position of the pulse in the fiber, so that the backscattered power for a given pulse wavelength is altered due to the changed interference condition. However, when compensating the larger distance between scatterers by also increasing the wavelength (decreasing the laser frequency v), the original interference condition and backscatter power is restored to a good approximation. Hence, by comparing a wavelength scan of the backscattered power I in the unstrained state and a wavelength shifted backscatter signature of a second scan in the strained state, one can quantitatively determine the local strain magnitude.

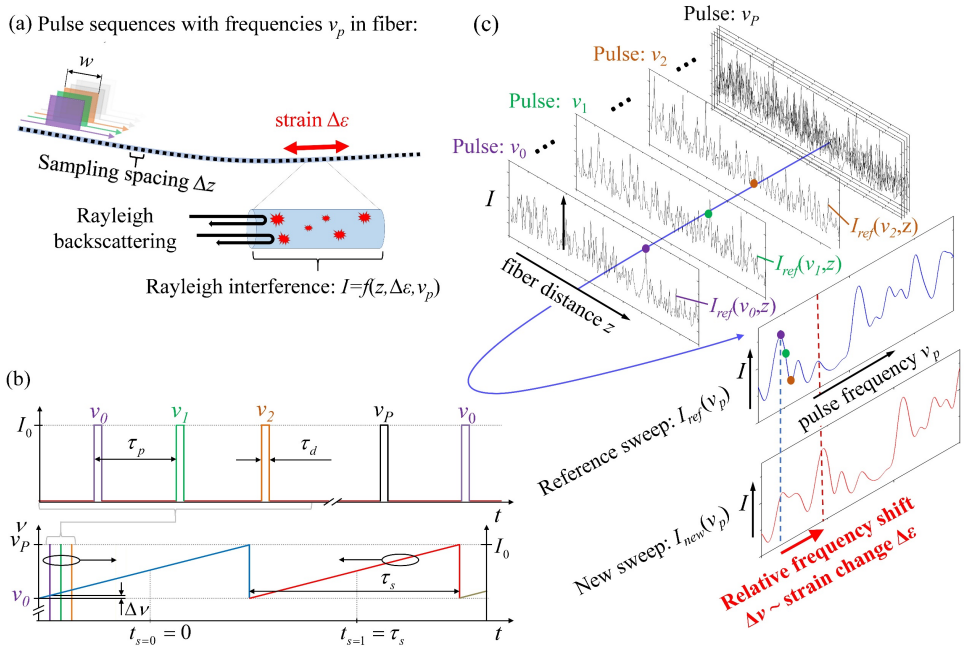

Fig. 1. (a) Schematic of strain change ($\Delta\varepsilon$) measurement along the fiber using Rayleigh backscatter interference from coherent pulses propagating along the fiber. (b) Schematic: sequential generation of pulses of different optical frequencies v_0 to v_P during periodic laser frequency sweep for wavelength-scanning of backscatter traces (pulse duration τ_d, pulse period τ_p, sweep period τ_s, pulse peak power I_0). (c) Schematic of backscatter trace evaluation at a given position z: relative frequency shift of the $I_{new}(v_p)$ vs $I_{ref}(v_p)$ backscatter power signature corresponds to the local strain change $\Delta\varepsilon$.

The backscatter signature shift along the frequency axis has previously been calculated by a least mean square (LMS) correlation approach [20]. In previous work, we derived a general expression for the distance-resolved and time-resolved coherent Rayleigh reflectometry for an optical pulse propagating along the fiber in [20]. For the generation of simulated ANN training data, a simplified model without the spatial and temporal dependencies is sufficient. The synthesis of ANN training data for the correlation task is presented in Section 2.2. The underlying physical model is described by the following parameters and assumptions:

Optics EXPRESS

The backscattered power from a single-mode fiber (SMF) can be modeled by a set of N discrete scatterers with a uniform statistical positional (distance z) distribution along the fiber and an independent uniform distribution of their reflectivities r. The detected Rayleigh backscattering power I from a monochromatic rectangular pulse of the peak power I_0 and the duration τ_d can be described as

$$I = I_d + I_{coh} = \sum_{i=1}^{N} r_i I_0 + 2\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sqrt{r_i r_j}\, I_0 \cos\left[2\pi\nu_0\tau_{ij}\right]\exp\left(-\left|\tau_{ij}\right|/\tau_c\right) \qquad (1)$$

where $\tau_{ij} = \tau_i - \tau_j = 2n_g z_{ij}/c_0$ are the scatterers' relative delays ($\tau_{ij} < \tau_d$), $z_{ij} = z_i - z_j$ are the physical separations between the individual scatterers, and ν_0 is the pulse optical frequency. The limited coherence of the light source is described by the degree of coherence function with τ_c as the coherence time. The detected power I consists of the direct summation power component I_d of the individual scatterer reflectivities r, and an interference component I_{coh} originating from coherent double summation of the scatterers' temporal phase differences. The spatial resolution of the sensor is defined by half the pulse width w in the fiber $w/2 = c_0\,\tau_d/(2n_g)$ where c_0 is the velocity of light and n_g is the fiber group refractive index. Polarization preservation within $w/2$ is assumed.

The direct summation component I_d is independent of relative scatterer displacements. However, the phase term of the interference component I_{coh} is a function of the temporal delay changes between the scatterers $\Delta\tau_{ij}$, as well as changes of the optical frequency ν_0 of the pulse. The delay change relation $\Delta\tau_{ij}/\tau_{ij}$ is proportional to the pulse optical frequency change relation $-\Delta\nu_p/\nu_0$ and related to strain changes $\Delta\varepsilon$ and temperature changes ΔT by the fiber's strain coefficient $K_\varepsilon \approx -0.78$ and temperature coefficient $K_T \approx -6.92\times10^{-6}$ [21] as

$$\begin{aligned}\frac{\Delta\nu_p}{\nu_0} &= -\frac{\Delta\tau_{ij}}{\tau_{ij}} = -\left(1-p_e\right)\Delta\varepsilon = K_\varepsilon\Delta\varepsilon \approx -0.78\Delta\varepsilon \\ \frac{\Delta\nu_p}{\nu_0} &= -\frac{\Delta\tau_{ij}}{\tau_{ij}} = -\left(\xi+\alpha\right)\Delta T = K_T\Delta T \approx -6.92\times10^{-6}\Delta T\end{aligned} \qquad (2)$$

where $p_e \approx 0.22$ is the effective strain-optic coefficient of silica, ξ is the thermo-optic coefficient of silica, α is the fiber thermal expansion coefficient, and the strain change is defined as $\Delta\varepsilon = \Delta z_{ij}/z_{ij}$. The equivalent impact of both, strain changes $\Delta\varepsilon$ and laser frequency changes $\Delta\nu_p$, on the detected power can thus be described as

$$I\left(\Delta\nu_p,\Delta\varepsilon\right) = I_d + 2\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sqrt{r_i r_j}\, I_0\cos\left\{2\pi\left(\nu_0+\Delta\nu_p\right)\tau_{ij}\left[1+K_\varepsilon\Delta\varepsilon\right]\right\}\exp\left(-\left|\tau_{ij}\right|/\tau_c\right). \qquad (3)$$

This means that the impact of strain changes on I can be compensated by equivalent optical frequency changes $\Delta\nu_p$ under the assumption that $\Delta\varepsilon$ is uniform within $w/2$.

The measurement task is to determine strain changes from the corresponding optical frequency changes. The fiber under test (FUT) is therefore periodically probed with a pulse repetition frequency $f_p = 1/\tau_p$ and pulse peak power I_0 while periodically tuning the laser frequency with a sweep rate $f_s = 1/\tau_s$, as indicated in Fig. 1(b). The backscatter traces from the individual pulses and pulse optical frequencies for each wavelength sweep are recorded and analyzed as depicted in Fig. 1(c). The pulse optical frequencies are indexed as $\nu_p = \nu_0 + \Delta\nu_p$ with discrete frequency steps of $\Delta\nu$ as $\Delta\nu_p = p\Delta\nu$, with $p = 0,1,2,\ldots,P$. $P+1$ backscatter traces are measured during each laser sweep period τ_s. As indicated in Fig. 1(c), frequency shift-equivalent strain changes $\Delta\varepsilon$ are obtained for each sampling position z in the fiber by computing the shift of new sweep results $I_{new}(\Delta\nu_p)$ relative to a reference sweep result $I_{ref}(\Delta\nu_p)$. The effective strain measurement repetition rate is $f_s = 1/\tau_s$.

Optics EXPRESS

2.2. Synthetic training data generation using Rayleigh backscatter model

The parameters and assumptions defined in Section 2.1 are used to simulate real measurement results of frequency sweeps at a given position z. K independent sets of N scatterers are synthesized with uniformly distributed reflectivities r ~unif(0,1), and uniformly distributed temporal delays τ within the pulse duration τ ~unif(0,τ_d). Each of the K simulated training sequences consists of a pair of simulated sweep results: one simulated sweep $I_{ref}(\Delta\nu_p)$ as reference with zero strain change ($\Delta\varepsilon = 0$) and one sweep result $I_\varepsilon(\Delta\nu_p,\Delta\varepsilon_{train})$ with a random training strain change $\Delta\varepsilon_{train}$:

$$I_{ref}\left(\Delta\nu_p\right)=\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sqrt{r_i r_j}\cos\left[2\pi\left(\nu_0+\Delta\nu_p\right)\tau_{ij}\right]\exp\left(-\left|\tau_{ij}\right|/\tau_c\right)\quad\text{for } p\in 0,\ldots,P$$

$$I_\varepsilon\left(\Delta\nu_p,\Delta\varepsilon_{train}\right)=\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sqrt{r_i r_j}\cos\left[2\pi\left(\nu_0+\Delta\nu_p\right)\tau_{ij}\left(1+K_\varepsilon\Delta\varepsilon_{train}\right)\right]\exp\left(-\left|\tau_{ij}\right|/\tau_c\right). \quad (4)$$

Linear factors and offset components from Eq. (3) are redundant and are omitted since z-score (standard score) normalization of each sequence is conducted before the training. A training data set C of K sweep pairs, as defined in Eq. (4), is generated for a defined strain training range of ± $\Delta\varepsilon_{max}$, with $\Delta\varepsilon_{train}$ from a uniform distribution $\Delta\varepsilon_{train}$ ~unif(-$\Delta\varepsilon_{max}$,$\Delta\varepsilon_{max}$). A single output ANN is trained to predict the label $\Delta\varepsilon_{train}$ from the simulated sweep pairs. The synthesized sweep results are combined in a matrix Q with dimension $K \times 2(P + 1)$, where the kth row contains the kth simulated pair of reference sweep and strain sweep (i.e., a concatenation of Eq. (4)) as

$$Q=\left(I_{ref}\left(\Delta\nu_0\right)^{\mathrm{T}},\ldots,I_{ref}\left(\Delta\nu_P\right)^{\mathrm{T}},I_\varepsilon\left(\Delta\nu_0,\Delta\varepsilon_{train}\right)^{\mathrm{T}},\ldots,I_\varepsilon\left(\Delta\nu_P,\Delta\varepsilon_{train}\right)^{\mathrm{T}}\right). \quad (5)$$

A noise matrix $Noise_\varepsilon$ of the same dimension as Q with elements sampled from a normal distribution with mean zero and standard deviation σ_ε is added to the ideal training matrix Q. The noise matrix is normalized by scalar multiplication with the approximate mean backscatter power $N/2$, which corresponds to the direct-summation power term I_d in Eq. (1) with $\bar{r}\approx 0.5$. Thus, the synthesized noisy training data matrix C is defined as

$$C = Q+\frac{N}{2}Noise_\varepsilon. \quad (6)$$

Adding noise to the idealized synthetic training data is crucial to avoid overfitting to the training data and make the trained network able to generalize for strain prediction from noisy measurement data. Z-score normalization of C is conducted for each sweep pair. $K = 4000000$ sweep pairs are used for the training in Section 4. The parameter σ_ε, which defines how much noise is added to the training data, is treated as a hyperparameter of the model (see Section 4).

2.3 Reference approach: Least mean square (LMS) correlation analysis

We previously proposed to use a least mean square (LMS) correlation approach to calculate the frequency shift [20]. This method already outperforms standard cross-correlation and is used as a reference to evaluate the ANN prediction performance throughout this paper. The measured strain change-equivalent frequency shift $\Delta\nu_m$ of a new sweep measurement $I_{new}(\Delta\nu_p)$ relative to a reference sweep $I_{ref}(\Delta\nu_p)$ corresponds to q^* that minimizes the LMS correlation result $R_{ref,new}(q)$ as

$$\Delta\nu_m = q^*\Delta\nu = \arg\min_q\left[R_{ref,new}\left(q\right)\right]\Delta\nu \quad (7)$$

where

Optics EXPRESS

$$R_{ref,new}(q)=\begin{cases}\dfrac{1}{m-q}\displaystyle\sum_{p=0}^{m-q}\left[I_{ref}\left(\Delta\nu_{p+q}\right)-I_{new}\left(\Delta\nu_{p}\right)\right]^2 & q\geq 0\\ R_{new,ref}(-q) & q<0\end{cases}\quad \text{with } |q|<m. \tag{8}$$

$\Delta\nu_m$ is quantized by the frequency step size $\Delta\nu$. The frequency shift resolution can, however, be improved by a cubic polynomial interpolation of $R_{ref,new}(q)$ (also compare Fig. 6(a)). Using Eq. (2), the high-resolution strain results $\Delta\varepsilon$ are calculated from $\Delta\nu_m$' of the interpolated results $R'_{ref,new}(q')$ as

$$\Delta\varepsilon=\frac{\Delta\nu_m'}{K_\varepsilon\nu_0}. \tag{9}$$

This method already proved more stable and less error-prone than a standard cross correlation analysis. This performance improvement in comparison to a cross-correlation, especially for a reduced number of input samples, has also been confirmed for a similar LMS analysis [33]. Although the LMS implementation is considerably faster than Pearson's correlation coefficient analysis used for similar applications [21,22], the computation time is still too high for real-time data processing. A major aim of using ANNs for strain prediction is to reduce the computation time for real-time applications.

3. Experimental setup

The robust and low-component cost experimental setup has been described in detail [19,20]. The direct-detection architecture is depicted in Fig. 2(a).

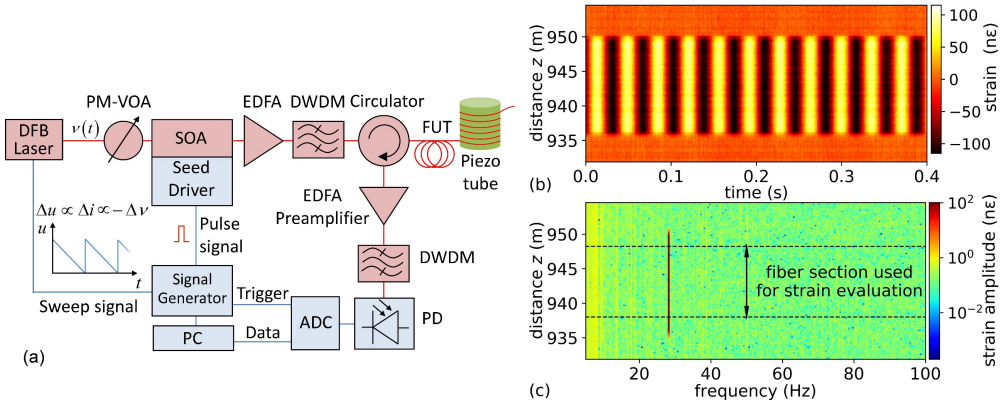

Fig. 2. (a) Schematic of the WS-COTDR implementation. (b) Example of ANN strain results along the fiber section wound around the piezo tube (from z ≈936.5 m to z ≈950.5 m). (c) FFT result of the same measurement (f_s = 1 kHz, 28 Hz signal, strain amplitude 100 nε, from 2 s measurement time) with indication of fiber section used for strain evaluation in Section 4.

A distributed feedback (DFB) laser diode (center wavelength of 1549.89 nm, ν_0 ≈193.36 THz) is used as an optical source in continuous wave (CW) mode. The laser linewidth of 1.3 MHz ensures adequate coherence within the pulse duration, which is typically 10 ns to 20 ns. A 120 MHz bandwidth signal generator supplies the rectangular signal for optical pulse shaping (with the pulse repetition frequency f_p and pulse duration τ_d). The pulse signal is fed via a seed driver to a semiconductor optical amplifier (SOA) with high extinction ratio (70 dB) and fast switching speed (1 ns). The signal generator also supplies the synchronized saw-tooth signal with the frequency f_s for direct current modulation of the laser. The resulting laser wavelength sweep due to the laser current modulation is, however, not ideally linear. The backscatter results of the non-linear sweeps must, therefore, be numerically corrected. This is achieved by an ANN subnet prior to the strain prediction and is described in Section 4.1. The optical power variation during the sweep is negligible [20]. The power level of the CW signal

Optics EXPRESS

into the SOA is adjusted by a polarization maintaining variable optical attenuator (PM-VOA). After the SOA, the optical pulses are amplified by an erbium-doped fiber amplifier (EDFA) and the amplified signal is filtered by a dense wavelength division multiplexing (DWDM) filter (0.2 nm bandwidth, flat-top) to remove amplified spontaneous emission (ASE) noise. The pulse sequences with pulse peak powers I_0 of a few hundred mW enter the fiber under test (FUT) via a polarization insensitive optical circulator. All backscattered light from the FUT is amplified by a preamplifier EDFA and is subsequently ASE-filtered before detection by a 125 MHz PIN photodetector (PD) with integrated transimpedance gain. Triggered by the signal generator, the backscatter signal is sampled at 500 MS/s and 14 bit resolution by an analog-to-digital converter (ADC) for further processing by a personal computer. The ADC sampling rate translates to a spatial sampling resolution of 20.4 cm. The maximum distance range z_{max} is limited by the round-trip time of the pulse as $z_{max} = c_0/(2n_g f_p)$. All measurements, except for the telecommunication cable in Section 4.4, have been conducted on a 970 m long FUT (Corning SMF-28e). A 14 m long section from z = 936.5 m to z = 950.5 m is wound around a piezo tube. An almost uniform strain distribution can be imposed onto this fiber section with a highly linear strain response as a function of the piezo's supply voltage. Figure 2(b) shows a strain measurement of the fiber section with the indication of the fiber section from z = 938.02 m to z = 948.24 m that is used to quantify all strain and performance parameters in Section 4.

4. ANN architecture, training and results

This section describes the ANN architecture, training, hyperparameter optimization and evaluates the ANN's performance against the LMS correlation approach. As mentioned in Section 3, the accurate strain prediction is achieved by a two-step process using a stacked ANN architecture that consists of two separately trained subnets with fully connected layers. The first ANN subnet corrects the not ideally linear laser frequency sweep. This ***linearization ANN*** step is crucial for a linear strain response of the sensor and is introduced in Section 4.1. The second subnet is the ***strain ANN*** which receives the output of the linearization ANN as input and predicts the strain for each frequency sweep. The sweep linearization and the strain prediction could be trained in a single integrated ANN. However, the model would have to generalize for a considerably more complex task. The complexity is reduced by training the two ANNs independently; this is possible since the labels for the linearization can also be synthesized. We therefore used a stacked architecture of the two ANNs trained individually for linearization and strain prediction. It is important to note that a single trained ANN can be used to compute strain changes for any fiber position and for any I_{ref} - I_{new} pair within $\pm \Delta\varepsilon_{max}$.

This section is structured as follows: After describing the linearization ANN in 4.1, the strain prediction ANN is covered in Section 4.2: Sensor-specific performance parameters that quantify the performance of a dynamic strain sensor are defined in Section 4.2.1, and the evaluation of the ANN (training, hyperparameter selection, and testing on real measurement data) is summarized in Section 4.2.2. In the following sub-sections, the ANN performance is compared to the previously used LMS correlation approach for the extremes of the sensor parameter range: Reduced signal-to-noise ratio (SNR) in Section 4.3.1, increased frequency step size $\Delta\nu$ in Section 4.3.2, and reduced length of correlation input (longer distance range) in Section 4.3.3. The most important improvement, namely the reduction of computation time, is summarized in Section 4.3.4. The real-time capability for arbitrary strain signals is demonstrated in Section 4.4 for the example of ground movement along a "dark fiber" in an urban environment.

4.1 Sweep linearization ANN

The actual laser frequency sweep is not ideally linear as it is defined in Eq. (3). The saw-tooth laser current modulation results in a swift frequency decrease followed by a monotonous but not constant increase in frequency. The few backscatter traces during decreasing frequency at

Optics EXPRESS

the beginning of each sweep period are discarded. The necessary correction (or: "linearization") of the measured $I_{meas}(\Delta\nu(p))$ prior to the application of the correlation algorithm has previously been computed by numerical interpolation [20]. For this calibration method, the actual frequency change $\Delta\nu(p)$ for each pulse during the sweep is experimentally obtained by phase unwrapping using a Mach-Zehnder interferometer. This is done once for each combination of laser sweep amplitude and sweep frequency. The computational sweep linearization step based on the calibration must then be conducted for each new sweep result and, so far, contributed significantly to the overall computation time

The computation time for the sweep linearization task is significantly reduced by using an ANN. This linearization ANN subnet is directly integrated with the strain ANN subnet, yielding a stacked model that conducts the linearization in a first step, and the strain prediction in subsequent layers (see model architecture in Fig. 3(a)). The training and hyperparameter optimization of the linearization ANN and the strain ANN are conducted independently.

Analogously to the training data generation of the strain ANN in Eqs. (4)-(5), the linearization ANN training data is simulated based on the Rayleigh backscatter interference model in Eq. (3). Pairs of M sweep result input vectors I_{meas_train} and M output vectors (labels) I_{linear_train} are simulated for the measured frequency change $\Delta\nu(p)$ as

$$I_{meas\_train}\left(\Delta\nu(p)\right)=\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sqrt{r_i r_j}\cos\left[2\pi\left(\nu_0+\Delta\nu(p)\right)\tau_{ij}\right]\exp\left(-\left|\tau_{ij}\right|/\tau_c\right)$$

$$I_{linear\_train}\left(\Delta\nu_p\right)=\sum_{i=2}^{N}\sum_{j=1}^{i-1}\sqrt{r_i r_j}\cos\left[2\pi\left(\nu_0+\Delta\nu_p\right)\tau_{ij}\right]\exp\left(-\left|\tau_{ij}\right|/\tau_c\right). \qquad (10)$$

The ANN is trained using the simulated (as measured) power results for each true frequency shift $I_{meas\_train}(\Delta\nu(p))$ as input vectors and the ideal (linearized) frequency sweep $I_{linear\_train}(\Delta\nu_p)$ with equal frequency steps $\Delta\nu$ as label vectors. Analogously to the generation of the training data for the strain ANN (see Eq. (4) and Eq. (6)), the linearization training data matrix is a superposition of ideal training data and a noise matrix $Noise_{lin}$ with elements generated from a Gaussian distribution $N(0,\sigma_{lin})$. Equivalent to Eq. (6), the linearization noise matrix is normalized by the mean backscattering power $N/2$. The model is trained using a variant of stochastic gradient descent with mean squared error (MSE) as loss function [34,35] and $M = 500000$ simulated sweep vectors.

A single hidden dense layer is sufficient to effectively conduct the sweep linearization. We found that adding $Noise_{lin}$ with $\sigma_{lin} = 0.1$ ensures robust prediction performance of the strain ANN for all considered measurement parameter values. Rectified linear units (ReLU) are used as nonlinear activation functions of the hidden layer. The specific number of nodes of the hidden layer $nodes_{lin}$ had only minor influence on the model's performance. We chose to set the number of nodes to $nodes_{lin} = 200$ for all measurement parameters in this paper.

Figure 3(b) shows the laser frequency shift $\Delta\nu(p)$ that has been obtained from the Mach-Zehnder calibration measurement as well as the ideal linear frequency change. An example of a single sweep measurement $I_{meas}(\Delta\nu(p),z)$ at the fiber section of the piezo is shown in Fig. 3(c). The linearized results $I_{linear}(\Delta\nu_p,z)$ predicted by the ANN are displayed in Fig. 3(d).

Optics EXPRESS

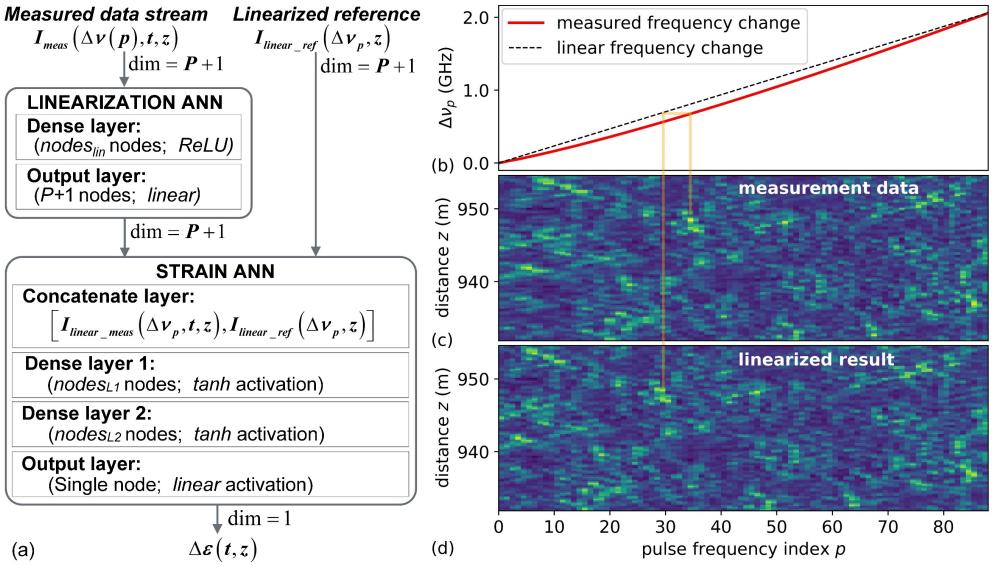

Fig. 3. (a) ANN architecture used for prediction from the raw measurement data stream: Stacked architecture of the linearization ANN and strain ANN. The arrows resemble the data flow during forward pass; data dimensions are added as dim. (b), (c), and (d) with common x-axis and indication of feature shift: (b) Calibration measurement of frequency change during the laser current modulation as a function of pulse sample number / frequency index p (f_s = 1 kHz, f_p = 100 kHz, τ_d = 20 ns, $\Delta\nu$ ≈23.8 MHz), and ideal (linear) frequency change with equal frequency steps $\Delta\nu$. (c) Measured $I_{meas}(\Delta\nu(p),z)$ for a single sweep around the piezo fiber section. (d) Linearized result $I_{linear}(\Delta\nu_p,z)$ obtained from the linearization ANN of the same measurement.

In comparison to the linearization interpolation, the linearization ANN speeds up the computation considerably by a factor of 272 for the specific conditions mentioned in Section 4.3.4. A computation time comparison for all subtasks is provided in Table 4. For real-time prediction, the pre-trained ANN weights (including the biases) are loaded into the integrated ANN for joint linearization and strain prediction, see Section 4.3.4. The schematic of the stacked architecture that is used for real-time strain prediction from not linearized frequency sweep inputs is shown in Fig. 3(a).

4.2 Strain prediction ANN

4.2.1 Performance parameters

The correlation ANN is trained on synthetic training data using MSE as loss function. However, to assess the model's performance on measurement data, other sensor performance parameters are more relevant: Namely the ***strain amplitude spectral density (ASD) noise*** and the ***total harmonic distortion (THD)***. The strain *ASD* noise characterizes the noise of a dynamic sensor system, and the *THD* is a measure for the linearity of the sensor response.

Both parameters are analyzed in the Fourier domain as indicated in Fig. 4(a). For our purposes we define the strain ***ASD* noise** as the mean strain amplitude of the strain spectral amplitude response over the frequency range from 300 Hz to 500 Hz (all measurements have been conducted with a sweep rate of 1 kHz). The ***THD*** is defined as the response to a sinusoidal strain excitation as strain amplitude relation of higher harmonic amplitudes up to the 10th harmonic $\Delta\hat{\varepsilon}_{2,\ldots,10}$ in comparison to the amplitude of the fundamental frequency $\Delta\hat{\varepsilon}_1$:

$$THD = \frac{\sqrt{\Delta\hat{\varepsilon}_2^{\,2} + \Delta\hat{\varepsilon}_3^{\,2} + \Delta\hat{\varepsilon}_4^{\,2} + \ldots + \Delta\hat{\varepsilon}_{10}^{\,2}}}{\Delta\hat{\varepsilon}_1}. \tag{11}$$

Optics EXPRESS

ASD noise and *THD* are calculated both from ANN strain predictions (ASD_{ANN} and THD_{ANN}) as well as from the LMS correlation results (ASD_{LMS} and THD_{LMS}). They are calculated as mean over 50 distance samples from z = 938.02 m to z = 948.24 m during a sinusoidal strain excitation of the piezo (see Fig. 2(b)).

4.2.2 ANN training and hyperparameter selection

After each training epoch (i.e., iteration over the training data), the *ASD* noise and *THD* are computed on measurement data to monitor the model's training progress. In addition, the mean squared loss on simulated data (loss computed on training data and validation loss computed on simulated data not used for training) is computed here to demonstrate the model's performance deviation on synthesized data (see Fig. 4(b)). The progression of *ASD* noise and *THD* on measurement data in comparison to training loss and validation loss on synthesized data during training is visible in Figs. 4(c) and 4(b), respectively. While validation loss on synthesized data is still improving, *ASD* noise and *THD* show signs of overfitting to the synthetic data. The performance relation of the ANN results compared to the reference approach (interpolation linearization + LMS correlation) is shown as $THD_{rel} = THD_{ANN}/THD_{LMS}$ as well as $ASD_{rel} = ASD_{ANN}/ASD_{LMS}$ after each training epoch.

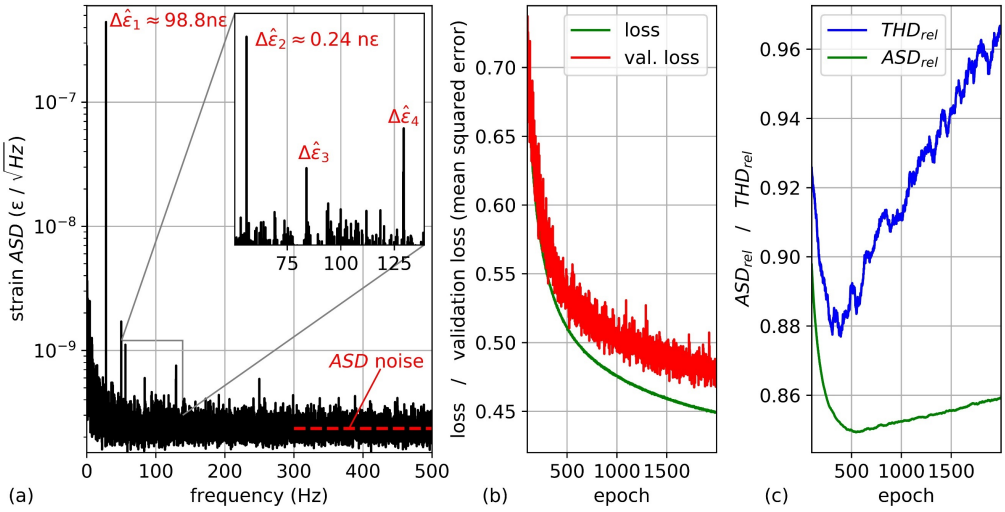

Fig. 4. (a) Strain *ASD* of a 28 Hz and 98.8 nε amplitude signal from ANN predictions at positions with strain modulation. The second harmonic suppression is 26.1 dB, or 52.2 dB in the more common power spectral density analysis in this field (mean of z = 938.02 m to z = 948.24 m, f_s = 1 kHz, f_p = 100 kHz, τ_d = 20 ns, measurement time = 20 s). Training progress of an exemplary ANN: (b) Loss and validation loss from synthesized training data, and (c) THD_{rel} and ASD_{rel} from measurement data ($nodes_{L1}$ = 600, $nodes_{L2}$ = 40, learning rate = 0.0001, $\Delta\varepsilon_{max}$ = ± 200 nε, ASD_{rel} and THD_{rel} filtered by 30 sample moving average).

The fact that the validation loss computed on synthesized data decreases monotonously whereas the *ASD* and *THD* parameters decrease after around 500 training epochs demonstrates the usefulness of using *ASD* noise and *THD* on measurement data as validation criteria: After a certain training time, the model overfits to the specific properties of the synthetic data. The reason for this discrepancy is mainly due to the difference of synthesized training data and real measurement data: We defined a specific optimization task based on a physical model of coherent Rayleigh scattering. However, an entirely correct representation of the real measurement data is not possible. Among other reasons, the discrepancies between measurement data and simulated data are due to several idealizations and assumptions that have been made when describing the physical principles and interactions in Eq. (4), e.g., rectangular pulse shape, polarization preservation within w/2. Other factors degrading the

Optics EXPRESS

correct emulation of measurement data are related to the actual signal generation and data acquisition: Sources of error are, for example, the limited photodetector and amplifier bandwidth, non-Gaussian noise components and pulse-to-pulse distortion due to pulse signal jitter, sampling jitter, and laser phase noise. This means that while the ANN is being trained on idealized synthetic data, the optimal sensor performance on actual measurement data is generally surpassed while the model's performance on the synthetic training data as well as on (also synthetically generated) validation data is still improving. This is clearly visible when comparing Figs. 4(c) and 4(b): While the MSE loss on synthetic validation data is still declining, the *ASD* noise and *THD* of predictions on measurement data is already degrading. Early stopping based on *ASD* noise and *THD* analysis is therefore applied. By monitoring the *ASD* and *THD* parameters on real measurement data during training and using these parameters for early stopping and hyperparameter selection, we address the domain adaptation problem from the synthetic training data (i.e., the source domain) to the real measurement data (the target domain) [36].

The reason for conducting *ASD* and *THD* analyses is that the measurement data is not labeled. The only reference would be the LMS correlation results which suffer from noise and may additionally comprise systematic errors and a nonlinear response. By validating *ASD* noise and *THD*, the learning progress can be monitored with more conclusive validation parameters with respect to real measurement data. Even though *ASD* noise and *THD* are good performance criteria, they are not suitable for directly defining a loss function to be used for training since the ANN has to generalize equally over the entire strain training range $\Delta\varepsilon_{train}$. Each training batch should ideally comprise a uniform label distribution within $\Delta\varepsilon_{train}$. This is, however, not ensured by a sinusoidal strain signal that is required for correct *THD* analysis.

Figures 4(b) and 4(c) clearly show that monitoring the strain noise and harmonic distortion during training on real measurement data is a better way to prevent overfitting than solely interpreting validation loss on synthetic data. The flow-chart in Fig. 5(a) depicts this training and evaluation approach which is used throughout this paper: All synthetic training data (i.e., each of the K = 4000000 input vectors of C) was normalized using z-scores. The label of the single-output ANN is the training strain $\Delta\varepsilon_{train}$.

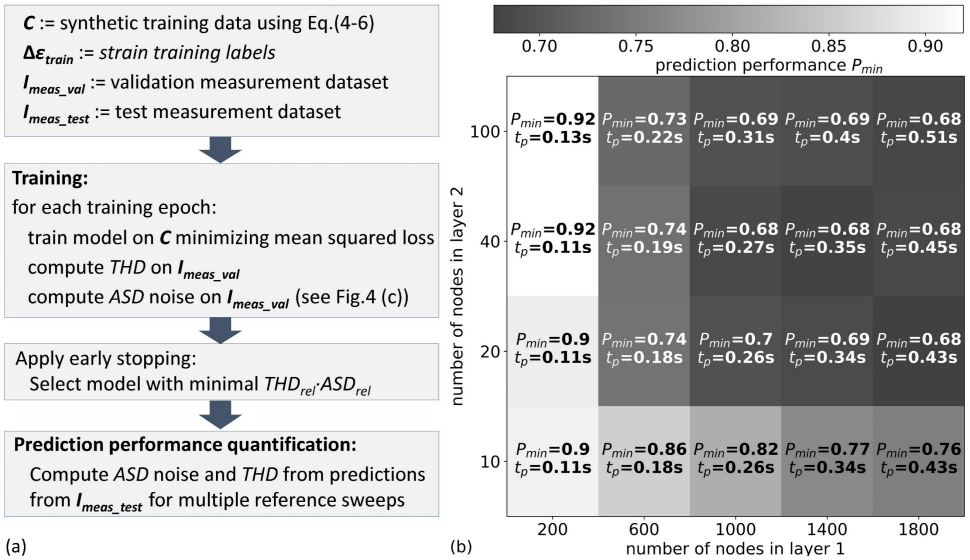

Fig. 5. (a) Schematic of strain ANN training, validation and test procedure. (b) ANN performance P_{min} for node combinations of layer 1 and layer 2 of a fully connected ANN, and indication of prediction time t_p per one million predictions (mean of P_{min} from six training sessions with randomly (Glorot uniform) initialized weights for each node combination; τ_d = 20 ns, lr = 0.0001, P = 85, $\Delta\varepsilon_{max}$ = ± 200 nε, N = 100, K = 4000000, batch size = 1024).

Optics EXPRESS

Two independent sets of measurement data are used: A validation measurement data set I_{meas_val} is used for hyperparameter optimization and to apply early stopping to avoid overfitting to the synthetic data. To this end, the training progress is monitored by computing THD_{rel} and ASD_{rel} after each training epoch, see Fig. 4(c). A test measurement data set I_{meas_test} is used to analyze the performance of the final model with optimized hyperparameters. The ANN's performance, defined as the product of THD_{rel} and ASD_{rel} predicted from I_{meas_val}, is calculated after each training epoch. For each training session, early stopping is applied to select the model with the best performance defined as P_{min} which is calculated as:

$$P_{\min} = \min_{epoch}\left(THD_{rel}\,ASD_{rel}\right). \tag{12}$$

Note that this is equivalent to simply minimizing the product of THD_{ANN} and ASD_{ANN}. THD_{rel} and ASD_{rel} are computed to visualize the ANN performance compared to the LMS correlation approach. The hyperparameter configuration of the P_{min} model is used to train the final model whose performance is analyzed on the test measurement data set I_{meas_test}. For testing, *ASD* noise and *THD* were calculated for 50 sensor positions as well as for a range of 10 different reference sweeps. Averaging the predictions for different sampling positions and for a range of reference sweeps yields a more conclusive comparison and performance result. These results are reported in Tables 1-3 for both: ANN predictions and LMS results. The *ASD* noise and *THD* analyses were conducted on the test measurement data set I_{meas_test} during a 100 nε sinusoidal excitation of the fiber by the piezo at the end of a 970 m long standard SMF (see Fig. 2). The modulation frequency was set to 20 Hz for I_{meas_val} and 28 Hz for I_{meas_test}. The size of the training data set C was $K = 4000000$ vectors, the size of the validation data set I_{meas_val} was 250000 vectors (5 s measurement), and the size of the test data set I_{meas_test} was 1000000 vectors (20 s measurement).

Various hyperparameters have been independently evaluated with this approach: the layer type (fully connected versus convolutional), the activation function of the hidden layers, number of hidden layers, number of nodes per layer, learning rate, and the training noise factor σ_ε. The Adam optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$) [34,35] was used for all training sessions. The ANNs were implemented in Tensorflow (v. 1.8.0) [37] using the Keras library (v. 2.2.0) [38]. A NVIDIA Quadro P4000 8GB RAM GPU was used for training and prediction in combination with a 16x Intel Xeon E5-1660 v3 3 GHz CPU running on Linux Open SuSe 15.0 64 Bit. The hyperparameter optimization was initially conducted for the same measurement settings as in [20]: $\tau_d = 20$ ns, $f_s = 1$ kHz, $f_p = 100$ kHz, $\Delta\nu = 11.94$ MHz. The selected hyperparameter configurations were then used as starting parameters to further optimize the hyperparameters for more extreme measurement settings (lower SNR, larger frequency step size $\Delta\nu$, and reduced number of sweep samples).

ANNs with fully connected hidden layers yielded a considerably better performance than convolutional ANNs. An added noise distribution with a standard deviation $\sigma_\varepsilon = 0.02$ (Eq. (6)) yielded optimal results for most parameter settings. In comparison to ReLU or sigmoid activation, using tanh as activation function yields both, a better performance of the trained model as well as shorter training time. A learning rate of $lr = 0.00003$ was suitable for all settings. Evaluating a range of combinations of the number of hidden layers and nodes per layer showed that two hidden layers are generally sufficient. Adding a third hidden layer did not improve the model's performance for the chosen measurement settings. The models' performances in terms of P_{min} are shown for a range of combinations of the number of nodes in the first hidden layer ($nodes_{L1}$) and the second hidden layer ($nodes_{L2}$) in Fig. 5(b). The output of the trained linearization ANN was used as input for the strain ANN training. In addition to P_{min}, the relevant prediction time per one million input sweep results t_p is given for each of the models. This analysis of node number combinations shows that model architectures with at least $nodes_{L1} = 1000$ nodes in the first hidden layer and $nodes_{L2} = 20$ in the second hidden layer yielded the best results. Further increasing the number of nodes did not significantly improve the performance but only increased the prediction time.

Optics EXPRESS

Further hyperparameter testing for the more extreme WS-COTDR measurement settings in Section 4.3.1 to Section 4.3.3 was therefore conducted with $nodes_{L1}$ = 1400 nodes in the first and $nodes_{L2}$ = 40 in the second hidden layer, tanh activation, and σ_ε = 0.02 training noise as starting values. It turned out that this initial hyperparameter configuration yielded the best results. Only the training data noise factor σ_ε had to be adapted for low-SNR measurement data in Section 4.3.1 and a coarser frequency step size in Section 4.3.2.

The ANN approach generally outperforms the LMS correlation approach in terms of *ASD* noise and *THD*. Moreover, the stacked ANN speeds up the linearization and strain computation considerably: compared to the LMS correlation approach, the computation time is decreased by a factor of 306, see Section 4.3.4 for details.

The strain of the raw measurement data stream was predicted by the stacked ANN architecture shown in Fig. 3(a) with the weights and biases of the independently trained linearization ANN and strain ANN. The linearized reference $I_{ref}(\Delta\nu_p,z)$ input sequence remained unchanged for each distance sample whereas the measurement data stream $I_{meas}(\Delta\nu(p),t,z)$ was normalized using z-scores before the concatenated arrays were passed on to Tensorflow.

4.3 Results

4.3.1 Performance at low-SNR measurement data

The hyperparameter selection was conducted for optimal measurement settings with the preamplifier EDFA current fully adjusted to the ADC input range without clipping. In real applications, however, optical loss along the fiber, for example due to connector degradation or mechanically induced macro-bend loss, may occur along the fiber and cannot be equally compensated by adapting pulse peak power or adjusting pre-amplification. SNR degradation due to fiber attenuation over long distances equally limits the maximum sensor length. A performance comparison for a range of signal-to-noise (SNR) levels has therefore been conducted. The SNR reduction is realized by decreasing the optical pulse peak power I_0 into the fiber without adjusting the preamplifier EDFA amplification. Table 1 summarizes the performance parameters for the LMS correlation approach and the ANN prediction, and compares the ANN performance with the LMS performance for a range of SNR-reduced measurement results.

Table 1. Performance of the ANN approach compared to the interpolation + LMS approach.

Performance parameter	I_0 reduction / SNR reduction	Interpolation / LMS result	ANN prediction	ANN performance improvement
THD [%]	0 dB	0.452	0.337	**25.4%**
ASD [nε/√Hz]		0.273	0.230	**15.8%**
THD [%]	3 dB	0.556	0.421	**24.3%**
ASD [nε/√Hz]		0.318	0.265	**16.7%**
THD [%]	6 dB	0.604	0.512	**15.2%**
ASD [nε/√Hz]		0.446	0.351	**21.3%**
THD [%]	10 dB	*	1.457	*
ASD [nε/√Hz]		*	0.721	*
THD [%]	13 dB	*	3.455	*
ASD [nε/√Hz]		*	1.443	*

*The LMS algorithm generally fails for ≥ 10 dB SNR reduction and is therefore not shown for comparison. (τ_d = 20 ns, f_s = 1 kHz, f_p = 100 kHz, P = 85, g = 0.24, $\Delta\varepsilon_{max}$ = ± 200 nε, lr = 0.00003, σ_ε = 0.02 for SNR reduction of 0 dB and 3 dB, σ_ε = 0.06 for SNR reduction of 6 dB, σ_ε = 0.4 for SNR reduction = 10 dB, and σ_ε = 0.8 for SNR reduction = 13 dB)

The same ANN was used for the performance analysis of optimal settings (0 dB loss) and the measurement SNR-reduced by 3 dB. However, ANNs that have been trained for high-SNR input predict a slightly reduced strain amplitude response when used on low-SNR measurement data (≥ 6 dB SNR reduction). This is due to a slight distortion of the prediction

Optics EXPRESS

accuracy toward the limits of the strain range of the training data $\Delta\varepsilon_{train}$ when the training noise (σ_ε) and the noise of the measurement data differ significantly from each other. This issue was addressed by training with an adapted noise factor σ_ε. Correct strain amplitudes are predicted for increased training noise of $\sigma_\varepsilon = 0.06$ for 6 dB, $\sigma_\varepsilon = 0.4$ for 10 dB, and $\sigma_\varepsilon = 0.8$ for 13 dB SNR reduction of measurement data. If considerable variations of the SNR along the sensor fiber are present, different SNR-specific ANNs would have to be used for accurate strain predictions.

Using ANNs, a significant performance improvement can be seen for the entire SNR range, especially for very low-SNR measurement data. It must be noted that the low-SNR test data from ≥ 10 dB SNR reduction could not be reliably processed using the LMS approach, whereas the ANN consistently predicts meaningful strain values even for 13 dB SNR reduction.

The optical loss budget of a DAS sensor system is a crucial performance parameter for a range of applications and also determines the maximum distance range of the sensor. The significant performance improvement using the ANN-based strain prediction approach for low backscatter signals will therefore extend the distance range compared to standard correlation-based methods.

4.3.2 Coarser frequency step size

It is a common problem that peak interpolation of severely under-sampled peaks, only a few samples wide, suffers from limited accuracy when determining peak positions or peak position shifts. Polynomial interpolation approaches of effectively under-sampled peaks, spectra, or, in this case, a (negative) LMS correlation peak (see Fig. 6(a)), lead to somewhat quantized peak position results, i.e., a nonlinear peak shift response.

This effect is evident in Fig. 6(b) for interpolated LMS correlation shift results from a measurement of an ideal triangular strain excitation that has been sampled with a coarser frequency step size of $\Delta\nu = 35.6$ MHz. We previously defined a conservative maximum limit for the frequency step size $\Delta\nu$ that still ensures correct interpolation results, i.e., a linear strain response [20]. This limit is a function of the pulse duration and is characterized for the LMS correlation approach by the factor $g = \Delta\nu\tau_d$ with $g < 0.25$. For small $g \leq 0.25$, high sensor linearity has been demonstrated over four orders of magnitude and down to peak amplitudes of 47.5 pε [20].

Increasing $\Delta\nu$, the number of sampling points that are available for correct minimum interpolation is effectively decreased, see Fig. 6(a). Therefore, the LMS correlation result after cubic polynomial interpolation exhibits a severely nonlinear strain response, as shown for $g = 0.71$ in Fig. 6(b). The "strain quantization effect" of the interpolated LMS results corresponds to the frequency step size $\Delta\nu$ and manifests itself in significant *THD* degradation.

Optics EXPRESS

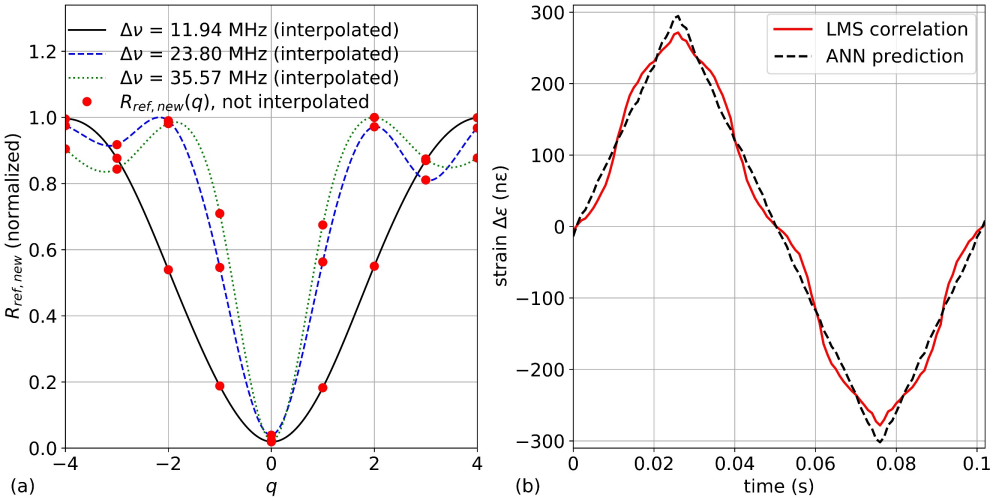

Fig. 6. (a) LMS correlation result $R_{ref,new}(q)$ from Eq. (8) for one distance sample, and cubic interpolation for: $g = 0.24$, $g = 0.48$ and $g = 0.71$ indicating reduced data quality for accurate minimum interpolation results. (b) Strain results during a 300 nε triangular modulation using the LMS approach and ANN predictions for coarser frequency step size of $\Delta\nu = 35.6$ MHz ($g = 0.71$, $f_s = 1$ kHz, $\tau_d = 20$ ns, $\Delta\varepsilon_{max} = \pm 500$ nε, mean of $\Delta\varepsilon$ from $z = 938.02$ m to $z = 948.24$ m).

Individual ANNs have been optimized, trained and tested for strain prediction on measurement data for a range of values of g: $g = 0.24$, $g = 0.48$, and $g = 0.71$. As mentioned above, the same ANN architecture and hyperparameters as in Section 4.2.2 yielded the best performance. Only the training noise σ_ε had to be increased for correct strain amplitude predictions for $g > 0.25$. The ANNs' performance was compared to the LMS correlation approach in Table 2 for ANNs trained for strain ranges of $\Delta\varepsilon_{max} = \pm 200$ nε and $\Delta\varepsilon_{max} = \pm 500$ nε, respectively.

Table 2. Performance of the interpolation + LMS results in comparison to the ANN predictions for $g = 0.24$ and coarser frequency step sampling with $g = 0.48$ and $g = 0.71$.

Parameter	Frequency step size $\Delta\nu$	Factor $g = \Delta\nu\tau_d$	Interpolation / LMS corr.	ANN prediction ($\Delta\varepsilon_{max} = \pm 200$ nε)	ANN prediction ($\Delta\varepsilon_{max} = \pm 500$ nε)	ANN improvement ($\Delta\varepsilon_{max} = \pm 200$ nε)
THD [%]	11.94 MHz	0.24	0.452	0.337	0.550	**25.4%**
ASD [nε/√Hz]			0.273	0.230	0.273	**15.8%**
THD [%]	23.80 MHz	0.48	5.748	0.594	0.792	**89.7%**
ASD [nε/√Hz]			0.275*	0.233	0.271	*
THD [%]	35.57 MHz	0.71	21.189	1.876	1.665	**91.1%**
ASD [nε/√Hz]			0.336*	0.314	0.349	*

* The value of ASD_{LMS} noise for $g > 0.25$ is not meaningful due to the strongly nonlinear response and is, therefore, not shown for comparison. Parameters: sinusoidal signal, 100 nε amplitude, $\tau_d = 20$ ns, $f_s = 1$ kHz, $f_p = 100$ kHz, $lr = 0.00003$, $\sigma_\varepsilon = 0.02$ for $g = 0.24$, $\sigma_\varepsilon = 0.04$ for $g = 0.48$, $\sigma_\varepsilon = 0.1$ for $g = 0.71$)

The ANN predictions exhibit considerably improved linearity of the strain response, which is clearly visible in Fig. 6(b). Whereas the ANN predictions closely follow the strain excitation, LMS correlation results exhibit reduced sensitivity around multiples of $\Delta\nu$. This means that strain cannot be correctly quantified with the LMS correlation approach for factors exceeding $g \approx 0.25$. Note, that this nonlinear sensor response is evident for even lower values of g when standard cross correlation analysis is used. It is also evident from Table 2 that ANNs trained for wider strain ranges (± 500 nε vs. ± 200 nε) exhibit a slightly reduced performance. This is expected since the ANN has to generalize for a wider shift range and, therefore, perform a more complex task.

In summary, ANNs clearly prove advantageous for coarser frequency step sampling, and potentially for under-sampled signals in general. This means that a wider strain range can be

Optics EXPRESS

scanned for a given number of frequency steps which effectively extends the possible strain measurement range when directly comparing two sweep results. We already demonstrated that the strain measurement range can be infinitely extended during a measurement without strain limitation by incorporating new sweep results into a continuously expanding reference $I_{ref,exp}(\Delta\nu_{p,exp})$ [20]. This method also promises further performance improvement but has not yet been implemented into the ANN prediction routine. Other methods to circumvent the limit of the strain prediction range are addressed in more detail in Section 5.

4.3.3 Reduced correlation input / extension of distance range

This subsection summarizes the ANN's performance for a reduced number of samples that are available as input for the LMS correlation analysis or the ANN, respectively. A reduced number of backscatter measurements per sweep may be necessary if a higher strain measurement repetition rate (sweep rate f_s) is required, and/or the distance range z_{max} is to be extended. The number of correlation input samples is inversely proportional to the product of sweep rate and distance range: $P \sim 1/(f_s\ z_{max})$. Naturally, fewer input data samples deteriorate the quality of the correlation.

To quantify this deteriorating effect, we compare the performance of the LMS and ANN approach for reduced correlation input size by means of reducing the pulse rate f_p and, therefore, extending the maximum distance range z_{max} while maintaining the strain repetition rate at f_s = 1 kHz. In addition to the z_{max} = 1020 m distance range (P = 85, results shown in Table 1), specific ANNs were trained, and their performance was analyzed for pulse repetition rate settings of: f_p = 50 kHz corresponding to z_{max} ≈2040 m (P = 42), and f_p = 20 kHz corresponding to z_{max} ≈5100 m (P = 16). The same hyperparameters as for P = 85 in Section 4.2.2 are suitable for the reduced input lengths. The prediction performance of models for 43 input samples per sweep and 17 input samples per sweep are summarized in Table 3.

Table 3. Performance comparison of the linearization interpolation + LMS correlation results and the ANN predictions for a reduced number of frequency samples per sweep for extended distance ranges z_{max}.

Parameter	Distance range z_{max}	Input samples for correlation	LMS correlation	ANN prediction	ANN performance improvement
THD [%]	2040 m	43	5.827	0.663	**88.6%**
ASD [nε/√Hz]			0.364	0.318	**12.6%**
THD [%]	5100 m	17	*	0.999	*
ASD [nε/√Hz]			*	0.501	*

* The LMS algorithm generally fails when only 17 pulse frequency samples are available. Hence, results are not shown for comparison. Constant sweep rate of f_s = 1 kHz for all measurements, $nodes_{lin}$ = 200, $nodes_{L1}$ = 1400, $nodes_{L2}$ = 40, σ_ε = 0.02, lr = 0.00003, f_s = 1 kHz, g = 0.48 for P = 42, g = 0.49 for P = 16, $\Delta\varepsilon_{max}$ = ± 200 nε, 100 nε strain amplitude.

It must be noted that the LMS correlation approach generally fails when only 17 input samples are provided. The ANN predictions, however, can be trusted even for significantly reduced input data. Hence, the use of ANNs allows for a reduction of the number of pulse frequencies per sweep. This considerably relaxes the limitation of the product of distance range and repetition rate of the WS-COTDR.

4.3.4 Computation time improvement for real-time measurement

The most important improvement is achieved in terms of computation time. Efficient ANN forward propagation for strain prediction from raw sweep data is computed on a GPU. Table 4 summarizes the computation times for the linearization task and the strain calculation tasks, as well as combined computation of both tasks using the reference approach (sweep interpolation / LMS correlation) and the ANNs.

Optics EXPRESS

Table 4. Computation time comparison of the reference approach (sweep linearization interpolation and LMS correlation) with the linearization ANN, the strain ANN, and the stacked ANN (see Fig. 3(a)).

Computation task	Computation time* (per one million input sweeps / input vectors)			ANN improvement (factor)
	Interpolation + LMS	Separate ANN subnets	Stacked ANN	
Linearization	33.21 s	0.122 s	-	**272.2**
Strain computation	94.77 s	0.353 s	-	**268.5**
Linearization + strain calculation	127.98 s	-	0.418 s	**306.2**

* GPU: NVIDIA Quadro P4000 8 GB RAM; Operating system: Linux, Open SuSe 15.0 64 bit; processor: 16 x Intel Xeon E5-1660 v3 3 GHz; backend: Tensorflow v. 1.8.0 implemented in Keras v. 2.2.0; ANN: $nodes_{lin} = 200$, $nodes_{L1} = 1400$, $nodes_{L2} = 40$, $P = 85$.

The computation of the linearization and the strain is 306 times faster when using a stacked ANN as compared to the LMS correlation approach with preceding interpolation. The stacked ANN architecture yields a speedup of more than 12% compared to calling the linearization ANN and the strain ANN separately.

Dynamic strain measurement at 2392 sensor positions could be conducted in real-time at a measurement repetition rate of 1 kHz with the current implementation. This real-time strain measurement capability can be further expanded to more than 10000 spatial sample locations by reducing the number of nodes of the hidden layers (compare prediction times in Fig. 5(b)), reducing the input sample size, and/or using a high-end GPU or several GPUs in parallel. Note that all necessary data preprocessing steps (reshaping the raw data and computing z-scores) can be conducted on CPU cores in parallel to the computation of the ANN forward pass on the GPU(s).

The capability of real-time prediction also solves the issue of mass data storage. The raw sweep data do not need to be stored since only the strain results from each sweep are of interest. Hence, the raw sweep data can be instantly discarded during the real-time ANN computation routine. This reduces the data volume by two to three orders of magnitude, depending on the number of pulse samples per sweep.

4.4 Demonstration of ground movement using a "dark fiber"

The potential of unused telecom fibers, or dark fibers, has been recently demonstrated, for example for geophysical applications [39–41] and rail track monitoring [42,43]. To demonstrate the applicability of the WS-COTDR for arbitrary signals in general, and the ANN approach in particular, we measured the strain distribution along a 1.3 km long dark telecom fiber between two locations of our institution. The telecom cable is buried at 0.8 m depth under a sidewalk in a protective duct at about 5 m distance in parallel to a car lane. The measured strain is in the axial direction along the fiber. Figure 7 shows two examples of urban near-surface strain distribution measurement showing quasi-static soil deformation and dynamic seismic wave propagation along the cable: Fig. 7(a) shows a measurement of the temporal strain distribution in which a car is pulling out of a parking space at $t \approx 2$ s, accelerates and drives in parallel to the fiber. The diagonal negative strain signature is caused by the deformation of the soil by the weight of the car accelerating to about 30 km/h (calculated from weight-induced strain signal along the fiber). The positive residual strain around $z = 647$ m is due to soil relaxation relative to the reference sweep at $t = 20$ ms after the car's weight is removed. Figure 7(b) shows the strain signature of a pedestrian walking along the buried cable. Each step causes visible surface wave propagation along the cable and temporary deformation of the cable due to the pedestrian's weight.

Optics EXPRESS

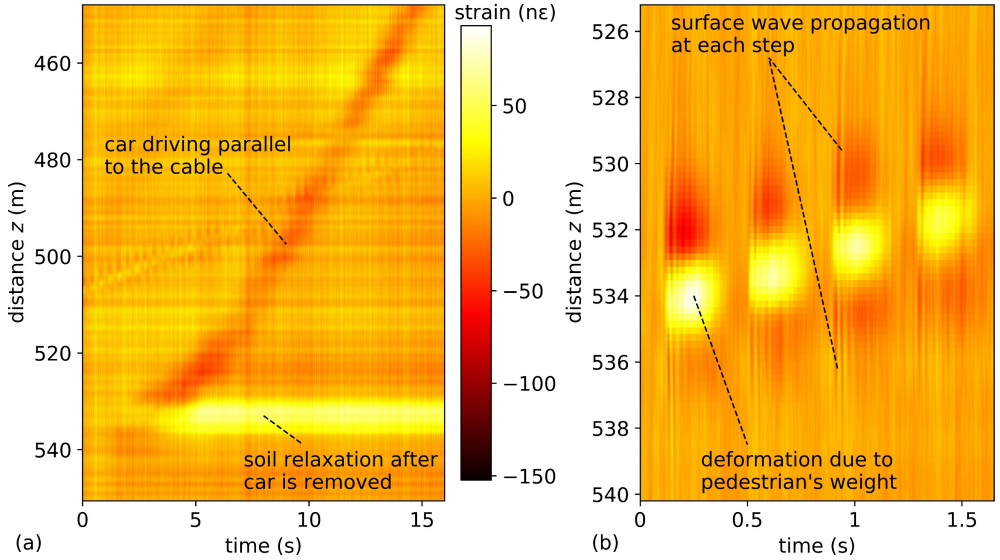

Fig. 7. Example for strain distribution measurement along a dark fiber in a telecom cable under a sidewalk in parallel to a road: (a) Car pulling out of a parking space and accelerating in parallel to the cable. (b) Pedestrian walking along the buried telecommunication cable. (a) and (b) with common strain scale; τ_d = 20 ns, f_s = 1 kHz, f_p = 100 kHz; ANN: $\Delta\varepsilon_{max}$ = ± 200 nε, $nodes_{lin}$ = 200, $nodes_{L1}$ = 1400, $nodes_{L2}$ = 40, σ_ε = 0.02, strain averaging of 5 samples along distance and time.

The subtle, almost vertical features are low-amplitude surface waves from vibration sources or physical impact sources at some distance from the fiber. They propagate over distances of several hundred meters and pass the fiber at various angles. This measurement on a dark fiber demonstrates that arbitrary strain signals and spatial strain distributions can be measured with real-time capability and nε-resolution from ANN predictions using the WS-COTDR approach.

5. Discussion

We showed that ANNs outperform conventional correlation algorithms with respect to all performance parameters for our application. The generation of synthetic training data is very well suited in cases where insufficient training data are available, or labels are either non-existing or error-prone. Artificial training data generation is particularly useful when the relation between ANN input and ANN output can be well-described by a physical or mathematical model. We used a physical model of the coherent interaction of the optical pulse with multiple discrete scatterers in the fiber for training data generation. However, an error-free representation of complex processes and their conversion to measurement data (ANN input data) is usually not possible. The associated issue of domain adaptation from the source domain (synthesized data) to the target domain (unlabeled measurement data) is addressed by applying early stopping during the training where the stopping criterion is defined in terms of the model's performance on measurement data with respect to the relevant sensor parameters: noise spectral density and linearity of the strain response. Also, the hyperparameters are selected using this domain adaptation approach. This method of ANN training and evaluation with combined synthetic data and measurement data is advantageous for a wide range of applications in science and technology, especially if the overall performance can be better described by application-specific parameters other than just the training labels.

Increasing the frequency shift range, and consequently the strain prediction range, is of interest for our application as well as for similar correlation tasks. The example of ANNs

Optics EXPRESS

trained for strain ranges of 400 nε and 1 με shows that ANNs trained for wider strain ranges somewhat deteriorate the model's performance. Improved performance for higher strain ranges can be achieved using various approaches: Several ANNs can be trained separately for smaller, but partly overlapping strain ranges. These strain range-specific sub-ANNs could then be integrated into a more complex ANN with an independently trained initial strain range classification subnet which selects the corresponding strain range sub-ANN for the final strain prediction. Another way to increase the accuracy of the strain prediction and extend the measurable strain range is to scan a much wider frequency range as a reference and adaptively select the matching reference sweep input range based on preceding predictions. This ensures a high strain measurement range while minimizing the number of sweep samples for a maximal product of sweep rate and distance range. A third option is to continuously update and extend the range of the reference scan with new sweep results, as demonstrated in [20]. We are confident that these more complex implementations will further improve the strain prediction accuracy.

We have demonstrated the real-world performance of the sensor using an underground telecommunication cable. Important application prospects using existing telecommunication infrastructure are the monitoring of railways and roads, active and passive seismic applications or ambient noise tomography [39], for example, for geothermal explorations. Other fields of application for this highly linear true strain sensor can be found in the structural health monitoring sector for damage assessment and lifetime prediction of critical infrastructures, such as bridges [20] and buildings, as well as the monitoring of pipes, pipelines and power cables.

With the results presented here, WS-COTDR can now compete in real-time measurements with phase resolved DAS, but at reduced maximal measurement repetition rates. A major advantage of the WS-COTDR approach is that the sweep results of several interrupted measurements can be compared to one another both on short and long-term timescales because each wavelength scan can be compared relative to a permanent and unambiguous reference scan. This contrasts with phase resolved DAS approaches, where the phase reference is usually lost after interruption of a measurement. Beyond the specific example presented here, ANN prediction-based signal shift analysis could also be used to improve the performance of various other correlation-based fiber optic sensor techniques in the time domain, frequency domain, and spatial domain [3,18,19,21,44–46]. The ANN interpolation and signal shift prediction approach can also be transferred to other fields in science such as spectroscopy, signal and time delay analysis, time-of-flight techniques, and pattern analysis in general.

6. Conclusion

Using ANNs for processing raw sensor data, we demonstrated substantial speed and accuracy improvements that enable real-time and high-resolution dynamic strain measurement along optical fibers using the WS-COTDR approach. Using fully connected ANNs, the performance of strain computation was considerably improved in comparison to correlation-based approaches. Two separate ANN subnets have been trained on synthesized training data for laser sweep linearization and strain prediction from relative signal shifts. The combined ANN not only performs the wavelength sweep correction and shift calculation task, but also provides high-accuracy interpolated results of the relative strain change. To moderate the issue of domain adaptation, the sensor-specific performance parameters noise spectral density and total harmonic distortion are used to evaluate the ANN's performance on measurement data during the training phase: Early stopping is applied to prevent overfitting of the synthesized training data. The same approach is used for hyperparameter optimization.

We showed that the ANNs outperform correlation approaches for low-SNR measurement data, for increased sampling step size, as well as for reduced number of correlation input samples. The greatest advantage, however, is the reduction of computation time by a factor of

Optics EXPRESS

306. This makes it for the first time possible to conduct real-time strain sensing using WS-COTDR. The real-time ANN strain prediction capability also considerably reduces the data storage requirements since raw data sweep measurements do not have to be saved for later computation. The applicability of the method for arbitrary strain distribution signals is demonstrated along a dark telecom fiber for dynamic and quasi-static ground movement in urban environment.

Funding

BAM Themenfeld-Projekt "Bewertung, Lebensdauerprognose und Instandsetzung von Brückenbauwerken" (BLEIB).

Acknowledgment

We thank all partners involved in the project.

References

1. A. H. Hartog, *An Introduction to Distributed Optical Fibre Sensors* (CRC Press, 2017).
2. L. Shiloh and A. Eyal, "Sinusoidal frequency scan OFDR with fast processing algorithm for distributed acoustic sensing," Opt. Express **25**(16), 19205–19215 (2017).
3. J. Li, J. Gan, Z. Zhang, X. Heng, C. Yang, Q. Qian, S. Xu, and Z. Yang, "High spatial resolution distributed fiber strain sensor based on phase-OFDR," Opt. Express **25**(22), 27913–27922 (2017).
4. D.-P. Zhou, L. Chen, and X. Bao, "Distributed dynamic strain measurement using optical frequency-domain reflectometry," Appl. Opt. **55**(24), 6735–6739 (2016).
5. A. Masoudi and T. P. Newson, "Contributed Review: Distributed optical fibre dynamic strain sensing," Rev. Sci. Instrum. **87**(1), 011501 (2016).
6. Y. Muanenda, "Recent advances in distributed acoustic sensing based on phase-sensitive optical time domain reflectometry," J. Sens. **2018**, 3897873 (2018).
7. S. V. Shatalin, V. N. Treschikov, and A. J. Rogers, "Interferometric optical time-domain reflectometry for distributed optical-fiber sensing," Appl. Opt. **37**(24), 5600–5604 (1998).
8. J. C. Juarez, E. W. Maier, K. N. Choi, and H. F. Taylor, "Distributed fiber-optic intrusion sensor system," J. Lightwave Technol. **23**(6), 2081–2087 (2005).
9. Z. Qin, T. Zhu, L. Chen, and X. Bao, "High sensitivity distributed vibration sensor based on polarization-maintaining configurations of phase-OTDR," IEEE Photonics Technol. Lett. **23**(15), 1091–1093 (2011).
10. Y. Muanenda, C. J. Oton, S. Faralli, and F. Di Pasquale, "A cost-effective distributed acoustic sensor using a commercial off-the-shelf DFB laser and direct detection phase-OTDR," IEEE Photonics J. **8**(1), 1–10 (2016).
11. Y. Dong, X. Chen, E. Liu, C. Fu, H. Zhang, and Z. Lu, "Quantitative measurement of dynamic nanostrain based on a phase-sensitive optical time domain reflectometer," Appl. Opt. **55**(28), 7810–7815 (2016).
12. G. Tu, X. Zhang, Y. Zhang, F. Zhu, L. Xia, and B. Nakarmi, "The development of an Φ-OTDR system for quantitative vibration measurement," IEEE Photonics Technol. Lett. **27**(12), 1349–1352 (2015).
13. Z. Wang, L. Zhang, S. Wang, N. Xue, F. Peng, M. Fan, W. Sun, X. Qian, J. Rao, and Y. Rao, "Coherent Φ-OTDR based on I/Q demodulation and homodyne detection," Opt. Express **24**(2), 853–858 (2016).
14. R. Posey, G. A. Johnson, and S. T. Vohra, "Strain sensing based on coherent Rayleigh scattering in an optical fibre," Electron. Lett. **36**(20), 1688–1689 (2000).
15. A. E. Alekseev, V. S. Vdovenko, B. G. Gorshkov, V. T. Potapov, and D. E. Simikin, "A phase-sensitive optical time-domain reflectometer with dual-pulse diverse frequency probe signal," Laser Phys. **25**(6), 065101 (2015).
16. A. Masoudi, M. Belal, and T. P. Newson, "A distributed optical fibre dynamic strain sensor based on phase-OTDR," Meas. Sci. Technol. **24**(8), 085204 (2013).
17. A. E. Alekseev, V. S. Vdovenko, B. G. Gorshkov, V. T. Potapov, I. A. Sergachev, and D. E. Simikin, "Phase-sensitive optical coherence reflectometer with differential phase-shift keying of probe pulses," Quantum Electron. **44**(10), 965–969 (2014).
18. J. Pastor-Graells, H. F. Martins, A. Garcia-Ruiz, S. Martin-Lopez, and M. Gonzalez-Herraez, "Single-shot distributed temperature and strain tracking using direct detection phase-sensitive OTDR with chirped pulses," Opt. Express **24**(12), 13121–13133 (2016).
19. S. Liehr, Y. S. Muanenda, S. Münzenberger, and K. Krebber, "Relative change measurement of physical quantities using dual-wavelength coherent OTDR," Opt. Express **25**(2), 720–729 (2017).
20. S. Liehr, S. Münzenberger, and K. Krebber, "Wavelength-scanning coherent OTDR for dynamic high strain resolution sensing," Opt. Express **26**(8), 10573–10588 (2018).
21. Y. Koyamada, M. Imahama, K. Kubota, and K. Hogari, "Fiber-optic distributed strain and temperature sensing with very high measurand resolution over long range using coherent OTDR," J. Lightwave Technol. **27**(9), 1142–1146 (2009).
22. L. Zhou, F. Wang, X. Wang, Y. Pan, Z. Sun, J. Hua, and X. Zhang, "Distributed strain and vibration sensing system based on phase-sensitive OTDR," IEEE Photonics Technol. Lett. **27**(17), 1884–1887 (2015).

Optics EXPRESS

23. C. M. Bishop, "Neural networks and their applications," Rev. Sci. Instrum. **65**(6), 1803–1832 (1994).
24. J. Schmidhuber, "Deep learning in neural networks: an overview," Neural Netw. **61**, 85–117 (2015).
25. M. Aktas, T. Akgun, M. U. Demircin, and D. Buyukaydin, "Deep learning based multi-threat classification for phase-OTDR fiber optic distributed acoustic sensing applications," Proc. SPIE **10208**, 102080G (2017).
26. L. Shiloh, A. Eyal, and R. Giryes, "Deep learning approach for processing fiber-optic DAS seismic data," in 26th International Conference on Optical Fiber Sensors (Optical Society of America, 2018), paper ThE22.
27. W. Zhaoyong, L. Luchuan, Z. Hanrong, L. Jiajing, W. Xiao, L. Bin, Y. Qing, C. Haiwen, and Q. Ronghui, "Smart distributed acoustics/vibration sensing with dual path network," in 26th International Conference on Optical Fiber Sensors (2018) (Optical Society of America, 2018), paper WF105.
28. J. Tejedor, J. Macias-Guarasa, H. F. Martins, J. Pastor-Graells, P. Corredera, and S. Martin-Lopez, "Machine learning methods for pipeline surveillance systems based on distributed acoustic sensing: A review," Appl. Sci. (Basel) **7**(8), 841 (2017).
29. A. K. Azad, L. Wang, N. Guo, H.-Y. Tam, and C. Lu, "Signal processing using artificial neural network for BOTDA sensor system," Opt. Express **24**(6), 6769–6782 (2016).
30. R. Ruiz-Lombera, A. Fuentes, L. Rodriguez-Cobo, J. M. Lopez-Higuera, and J. Mirapeix, "Simultaneous temperature and strain discrimination in a conventional BOTDA via artificial neural networks," J. Lightwave Technol. **36**(11), 2114–2121 (2018).
31. R. Ruiz-Lombera, J. M. Serrano, and J. M. Lopez-Higuera, "Automatic strain detection in a Brillouin optical time domain sensor using principal component analysis and artificial neural networks," in Proc. 2014 IEEE SENSORS (2014), pp. 1539–1542.
32. H. Wu, C. Zhao, R. Liao, Y. Chang, and M. Tang, "Performance enhancement of ROTDR using deep convolutional neural networks," in 26th International Conference on Optical Fiber Sensors (2018) (Optical Society of America, 2018), paper TuE16.
33. L. Zhang, Z. Yang, F. Gyger, M. A. Soto, and L. Thévenaz, "Rayleigh-based distributed optical fiber sensing using least mean square similarity," in 26th International Conference on Optical Fiber Sensors (Optical Society of America, 2018), paper ThE29.
34. D. P. Kingma and L. J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 [cs.LG] (2015).
35. S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of Adam and beyond," in Proceedings of International Conference on Learning Representations (2018).
36. S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," Mach. Learn. **79**(1-2), 151–175 (2010).
37. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: A system for large-scale machine learning," arXiv:1605.08695 [cs.DC] (2016).
38. "Keras Documentation," https://keras.io/.
39. P. Jousset, T. Reinsch, T. Ryberg, H. Blanck, A. Clarke, R. Aghayev, G. P. Hersir, J. Henninges, M. Weber, and C. M. Krawczyk, "Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural features," Nat. Commun. **9**(1), 2509 (2018).
40. S. Dou, N. Lindsey, A. M. Wagner, T. M. Daley, B. Freifeld, M. Robertson, J. Peterson, C. Ulrich, E. R. Martin, and J. B. Ajo-Franklin, "Distributed acoustic sensing for seismic monitoring of the near surface: A traffic-noise interferometry case study," Sci. Rep. **7**(1), 11620 (2017).
41. N. J. Lindsey, E. R. Martin, D. S. Dreger, B. Freifeld, S. Cole, S. R. James, B. L. Biondi, and J. B. Ajo□ Franklin, "Fiber-optic network observations of earthquake wavefields," Geophys. Res. Lett. **44**(23), 11792–11799 (2017).
42. W. Lienhart, C. Wiesmeyr, R. Wagner, F. Klug, M. Litzenberger, and D. Maicz, "Condition monitoring of railway tracks and vehicles using fibre optic sensing techniques," in Proc. Int. Conf. on Smart Infrastructure and Construction (ICE Publishing, 2016), pp. 45–50.
43. G. Cedilnik, R. Hunt, and G. Lees, "Advances in train and rail monitoring with DAS," in 26th International Conference on Optical Fiber Sensors (Optical Society of America, 2018), paper ThE35.
44. M. Froggatt and J. Moore, "High-spatial-resolution distributed strain measurement in optical fiber with rayleigh scatter," Appl. Opt. **37**(10), 1735–1740 (1998).
45. S. Liehr, M. Wendt, and K. Krebber, "Distributed strain measurement in perfluorinated polymer optical fibres using optical frequency domain reflectometry," Meas. Sci. Technol. **21**(9), 094023 (2010).
46. M. A. Soto, X. Lu, H. F. Martins, M. Gonzalez-Herraez, and L. Thévenaz, "Distributed phase birefringence measurements based on polarization correlation in phase-sensitive optical time-domain reflectometers," Opt. Express **23**(19), 24923–24936 (2015).