Deep Learning Enhanced Time-Domain Microwave Photonic Sensor

Xiaoyi Tian, Yiming Yan, Yeming Chen, Liwei Li, *Member, IEEE*, Luping Zhou, *Senior Member, IEEE*, Linh Nguyen, Robert Minasian, *Life Fellow, IEEE*, and Xiaoke Yi, *Member, IEEE*

***Abstract*—In this article, a new deep learning enhanced microwave photonic (MWP) sensor based on optical sideband processing with linear frequency modulated (LFM) pulse to enable fast and accurate sensing in the presence of noise and interference, is proposed and demonstrated. The LFM pulse produces a fast sweep of the optical sideband across the resonance, causing an envelope dip in the transmitted LFM pulse which enables fast interrogation. By optimizing the optical modulator bias point, the depth of this envelope dip can be increased, thus enhancing the interrogation resolution and facilitating higher accuracy. To enhance the noise resistance for practical uses and avoid adding unnecessary system complexity, we adopt a deep neural network (DNN) to process the raw interrogation output of this time-domain MWP sensor. The DNN can compress input data into a lower-dimensional representation and be trained to accurately estimate the target measurand despite noise and interference. Beside presenting a detailed theoretical model, as a proof-of-concept, we demonstrate the proposed scheme with a convolutional neural network (CNN) in measuring the glycerol solution concentration amidst thermal interference and system noise. The interrogation speed reached 2.5 MHz. Compared to the conventional MWP interrogation without bias control, using optimized bias voltages improves the accuracy of the estimation model based on the linear fitting of the manually extracted envelope dip by 2-fold. The CNN based prediction model achieves a root-mean-square error of 0.05%, which demonstrates an overall 4-fold higher accuracy than that of the reference model using traditional interrogation and filtering methods.**

***Index Terms*—Deep learning, deep neural networks, machine learning, microresonators, microwave photonics, photonic signal processing, sensors.**

I. Introduction

OPTICAL microresonators (OMRs), such as microrings, microdisks, and microspheres, are widely recognized as an ideal platform for demanding sensing applications thanks to their ability to strongly enhance light-matter interaction. Compared to traditional sensors, OMR sensors offer many advantages, including superior environmental sensitivity, a compact footprint, label-free detection capability, and compatibility with diverse functional materials and integration techniques, and hence have found applications in numeral fields ranging from detecting various physical fields, such as temperature, humidity, magnetic field, and pressures, to measuring various physical matters, including gases, solutions, biochemical molecules, and nanoparticles [1], [2], [3]. To improve the interrogation speed and resolution, different microwave photonic (MWP) technique based new interrogation schemes for OMRs have recently been demonstrated. According to the operation domain, these schemes to date can be categorized into two types. One transforms the spectral measurement from the optical domain to the radio frequency (RF) domain where a higher frequency resolution is readily available with the RF instrument. By adopting the OMR sensor in building up a MWP filter, the shift of the optical resonance can be transformed into the shift of an ultra-sharp RF spectral notch [4], [5] or an ultra-narrow RF signal [6], which further improves the interrogation resolution but has no benefit to the interrogation speed. The other acquires the interrogation output in the time domain to achieve a fast interrogation speed. By applying a linear frequency modulated (LFM) pulse for generating the interrogation light in a phase modulator, the optical resonance shift is interrogated as the shift of the envelope profile of the transmitted LFM pulse [7], so that the interrogation speed can be increased to be as fast as the pulse repetition rate. However, the sensing resolution of this approach requires a high extinction ratio and narrow linewidth of the OMR, which consequently leads to an increase in the design and fabrication complexity. Recently, we proposed a sensing scheme for OMRs by combining LFM pulse with MWP sideband processing [8]. Despite its potential to achieve high performance with reduced dependence on OMR parameters, this scheme has only been briefly analyzed in a simple model and implemented in controlled experimental conditions to sense temperatures, using a single specific pulse feature that was manually extracted. In fact, noise and interference that can affect the sensor are inevitable and can vary in different real-world applications. The ability to automatically and effectively distinguish the target measurand (TGT) from various interferences in a noisy environment is vital for achieving fast, accurate, and reliable sensing. It is, therefore, necessary and imperative to develop new techniques which can take full advantage of the

Manuscript received 31 March 2023; revised 22 June 2023; accepted 29 June 2023. Date of publication 4 July 2023; date of current version 2 December 2023. This work was supported by Australian Research Council. *(Corresponding author: Xiaoke Yi.)*

Xiaoyi Tian, Yiming Yan, Yeming Chen, Liwei Li, and Xiaoke Yi are with the School of Electrical and Information Engineering, the University of Sydney, Sydney, NSW 2006, Australia, and also with the University of Sydney Nano Institute (Sydney Nano), the University of Sydney, Sydney, NSW 2006, Australia (e-mail: xiaoyi.tian@sydney.edu.au; yiming.yan@sydney.edu.au; yche2958@uni.sydney.edu.au; liwei.li@sydney.edu.au; xiaoke.yi@sydney.edu.au).

Luping Zhou, Linh Nguyen, and Robert Minasian are with the School of Electrical and Information Engineering, the University of Sydney, Sydney, NSW 2006, Australia (e-mail: luping.zhou@sydney.edu.au; linh.n@sydney.edu.au; robert.minasian@sydney.edu.au).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2023.3292261.

Digital Object Identifier 10.1109/JLT.2023.3292261

© 2023 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

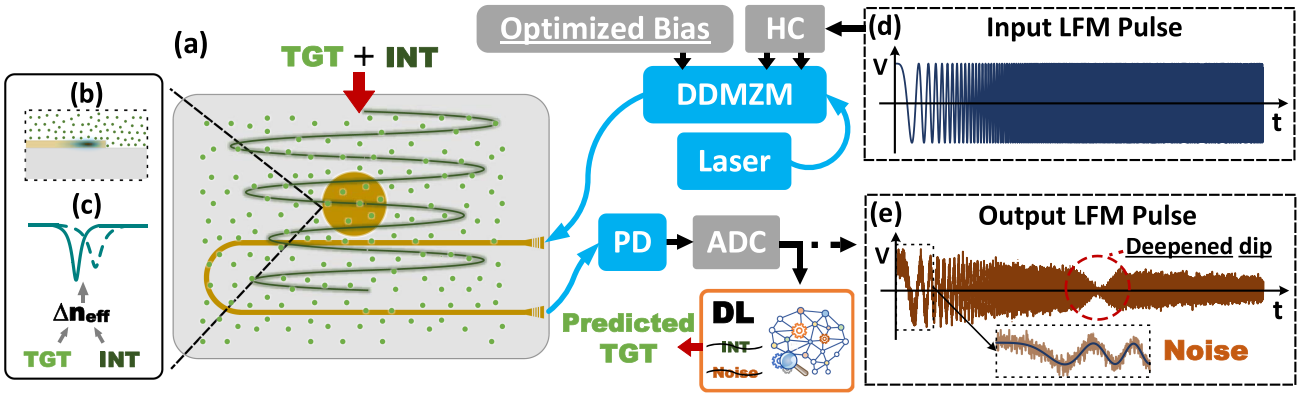

Fig. 1. (a) Schematic diagram of the proposed DL assisted noise resistant MWP sensing scheme; (b) evanescent waves of the simulated optical resonant mode field penetrating into the top cladding of microresonator; (c) changes of the optical resonance in response to the TGT and unwanted interferences (INT); (d) input LFM pulse (voltage (V) vs. time (t) plot) which works as the interrogation signals and provides fast interrogation speed; (e) output LFM pulse, i.e., the interrogation output, which is contaminated by system noise and contains a resonance-dependent deep envelope dip created by optimized DDMZM bias voltages. DDMZM: Dual-drive Mach-Zehnder modulator; HC: Hybrid coupler; PS: Power supply; PD: Photodetector; ADC: Analog-to-digital converter; DL: Deep learning.

time-domain sensing output to enhance the noise resistance of MWP sensors for practical uses.

Traditional denoising methods, such as Wiener filtering, exhibit limitations in adaptability and may result in the loss of crucial raw signal details [9]. Besides, these methods often fail to effectively remove interference that are highly correlated with the target signal. In recent years, deep learning (DL), as a subset of machine learning that involves training deep neural networks (DNNs), has emerged as an effective tool for enhancing the capabilities of sensors. Compared to the traditional data processing that relies on handcrafted features, DNNs, particularly the convolutional neural networks (CNNs), can effectively capture important patterns and relationships that are robust to noise and interference in the input data, and learn complex relationships between inputs and outputs to make accurate predictions simultaneously [10], [11]. Despite the fact that DNNs have gained prominence in speech recognition and general filtering compared to traditional methods such as Kalman and Wiener filters [12], [13], the utilization of DNN for time-domain MWP sensing applications has yet to be explored.

In this work, by integrating the DNNs into the time-domain MWP sensor based on optical sideband processing with the LFM pulse, we achieve a fast MWP sensing scheme that can operate effectively in the presence of noise and interference. A detailed theoretical model is presented to analyze the operational principle of the time-domain MWP interrogation scheme. By controlling the DC bias voltage of the optical modulator to optimize the power and phase of the optical sidebands that rapidly sweep through the optical resonance, the spectral profile of the optical resonance can be continuously transformed into a deep notch in the LFM pulse envelope in every short period. This enables fast and high-resolution interrogation of the microresonator sensor. To enable the MWP sensor to operate in the presence of noise and interference without increasing unnecessary system complexity, a DNN is adopted to process the raw interrogation output. The DNN can map the input to a compressed representation that consists of underlying structures and relationships that are invariant to the noise and interference, and subsequently makes accurate predictions of the TGT through training. As a proof-of-concept, the proposed scheme was implemented with a CNN based model for measuring glycerol solution concentration amid thermal interference and system noise. The utilization of short LFM pulses and optimized optical modulator bias voltages enables efficient interrogation of the resonance at a rapid rate of 2.5 MHz. The linear fitting prediction model established by using the envelope dip locations obtained by using the optimized bias voltages demonstrates a 2-fold increase in accuracy compared to that without bias control. By using the raw interrogation output to train a CNN, the established CNN-based prediction model achieves 4 times improvement in accuracy, as indicated by its root-mean-square error (RMSE) of around 0.05% in comparison with the RMSE of 0.19% of the linear model that uses the envelope dip locations under fixed bias conditions. This outcome underscores the power and potential of DL techniques in MWP sensing.

II. Principle and Methods

Fig. 1(a) shows the schematic diagram of the proposed DNN-enhanced MWP sensing scheme based on optical sideband processing with the LFM pulse. Before being injected into the on-chip microresonator sensor, the output of the laser is first modulated by a broadband LFM pulse in a dual-drive Mach-Zehnder modulator (DDMZM), creating two optical sidebands sweeping outward from the carrier at a fast rate, with the LFM pulse duration as the period. When an optical sideband sweeps through an optical resonance, its optical power and phase vary according to the optical power and phase transmissions of the resonance. This causes a resonance-dependent amplitude drop in the transmitted LFM pulse after photodetection. By optimizing the bias point of the DDMZM, the amplitude drop can be deepened to nearly zero, resulting in a more evident envelope dip carrying the response of the microresonator sensor. However, this response is often contaminated with external interference and system noise, making it difficult to accurately measure the TGT. To enhance the noise immunity of the MWP sensor without increasing hardware complexity, a DNN is proposed to process the raw interrogation output. The DNN filters out unwanted noise and interference and generates an accurate estimation of the TGT.

A. Principle of Operation

For a single optical resonant mode in an all-pass optical microresonator, which has a self-coupling coefficient of r and a single-pass amplitude transmission of a that reflects the transmission loss in the cavity, its optical transmission, $H(\lambda)$, with respect to the wavelength, λ, around the resonance wavelength, λ_{res}, can be expressed as [14]

$$H(\lambda) = \frac{r - ae^{i2\pi b}}{1 - rae^{i2\pi b}} \tag{1}$$

where $b = \frac{\lambda_{res}}{\lambda}$. At the resonance wavelength, λ_{res}, which is determined by

$$\lambda_{res} = \frac{2\pi}{N} n_{eff} L \tag{2}$$

where N is a positive integer, L is the physical transmission length of the resonant mode in the cavity, and n_{eff} is the effective index, the optical transmission exhibits a notch shape which has the depth or extinction ratio (ER) defined as

$$ER = \frac{(a+r)^2(1-ar)^2}{(a-r)^2(1+ar)^2} \tag{3}$$

and the width or full width at half maximum (FWHM) given by

$$FWHM = \frac{(1-ra)\lambda_{res}^2}{\pi n_g L\sqrt{ra}} \tag{4}$$

where n_g is the group index. At the wavelengths away from the resonance wavelengths, $H(\lambda)$ is equal to 1. As Fig. 1(b) shows, the optical field of the resonant mode not only distributes within the cavity but also extends beyond its surface as evanescent waves, allowing for intense light-matter interactions between the optical resonant mode and the claddings. Therefore, any environmental factors that can cause a variation in the refractive index of the cladding or cavity will perturb the optical field distribution and lead to an effective index change, Δn_{eff}, regardless of the TGT or unwanted interference. Based on (1)–(4), as illustrated in Fig. 1(c), the effective index change then leads to a variation of the spectral profile of the resonance transmission notch, which includes a resonance wavelength shift as well as the changes in the notch depth and width. The sensing operation is based on the interrogation and analysis of those resonance responses via DNNs.

The proposed sensing scheme conducts optical sideband processing with the LFM pulse to achieve fast and high-resolution interrogation of the resonant mode. Fig. 1(d) shows the temporal waveforms of an ideal LFM pulse. The LFM pulse has a constant amplitude and a repetition rate, v. In each repetition period, the frequency of the LFM pulse linearly varies over time, rising from zero to a maximum frequency at a constant rate of β. When the optical carrier wavelength, λ_C, is set close to the selected resonant mode of the microresonator sensor on the longer wavelength side and is modulated by the LFM pulse in the DDMZM, the optical sidebands will be created around the carrier and sweep outward from the carrier at the same speed of β with a period of $1/v$, and only the lower sideband (LSB) will travel across the selected resonance. Concerning only the optical carrier, LSB, and upper sideband (USB), the instantaneous intensity, $I(t)$, of the transmitted LFM pulse signals at the output of the photodetector (PD) in Fig. 1(a) at the time t can be approximated as [8]

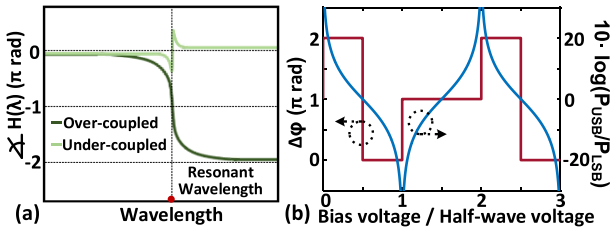

Fig. 2. (a) The phase transmission profile of over-coupled and under-coupled resonant mode of microresonators, where the red dot shows the resonant wavelength. (b) The power and phase relationship of the optical sidebands at different bias points of the DDMZM.

$$I(t) = 0.5\Re^2 P_C \left\{ P_{USB} \left| H\left(\frac{c\lambda_C}{c+\lambda_C\beta t}\right)\right|^2 + P_{LSB} + 2\sqrt{P_{LSB}P_{USB}\left| H\left(\frac{c\lambda_C}{c+\lambda_C\beta t}\right)\right|^2} \times \cos\left(\Delta\varphi + \angle H\left(\frac{c\lambda_C}{c+\lambda_C\beta t}\right)\right)\right\} \tag{5}$$

where $\Re$ is the PD responsivity, c is the speed of light, P_C, P_{USB}, and P_{LSB} are the optical power of the carrier, USB, and LSB, respectively, $\Delta\varphi$ is the phase difference between the RF components generated by the optical carrier beating with the USB and LSB, respectively. The resonance transmission is thus transformed to the envelope of the transmitted LFM pulse, where λ_{res} corresponds to the temporal location, t_d, which is equal to

$$t_d = \frac{c(\lambda_C - \lambda_{res})}{\lambda_C\lambda_{res}\beta} \tag{6}$$

Assuming that the DDMZM is biased close to the quadrature point, where $\cos\Delta\varphi = 1$, and the LFM pulse has an ideal transmission, the normalized intensity of the transmitted LFM pulse at t_d, $I'(t_d)$, can be calculated as

$$I'(t_d) = \frac{(1+K)^2}{4} \tag{7}$$

where K is equal to $-|H(\lambda_{res})|$ and $|H(\lambda_{res})|$ when the selected resonant mode is over-coupled and under-coupled, respectively. The resonance transmission notch shape is thus mapped to the transmitted LFM pulse envelope dip, while the intensity dip depth can only be around 6 dB. Fig. 2 depicts the phase transmissions of over-coupled and under-coupled resonant modes of optical microresonators and the DDMZM modulation characteristics. In the case of an over-coupled resonance, the phase transmission undergoes a sudden shift of 2π rad within the resonance region, and consistently changes by π rad at the resonance wavelength. In contrast, an under-coupled resonance also shows an abrupt phase change around the resonance wavelength, but the phase transmission at this point is always 0 rad. Besides, by

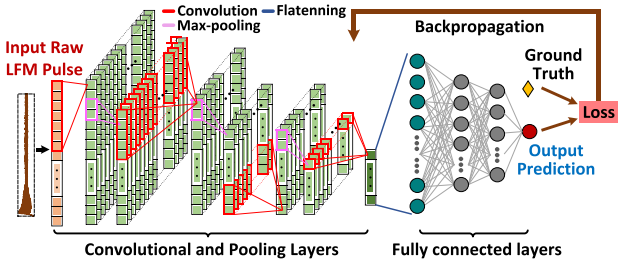

Fig. 3. The diagram of the CNN network structure used in this work.

tuning the bias voltage, the power ratio of USB and LSB can be adjusted over a wide range, while the phase difference $\Delta\varphi$ only changes in integer multiples of π radians. Therefore, there always exists an optimal bias point that can satisfy the following conditions

$$\cos\left(\Delta\varphi + \measuredangle H\left(\lambda_{res}\right)\right) = -1 \qquad (8-1)$$

$$P_{USB}|H\left(\lambda_{res}\right)|^2 = P_{LSB} \qquad (8-2)$$

which lead to $I'\left(t_d\right) = 0$ and an ER dependent maximum of $I(t)$. By adopting the optimal bias voltage during the MWP interrogation, the response of the resonant mode, which can be at arbitrary coupling states, can be interrogated with highly improved resolution as the changes of the transmitted LFM pulse envelope profile, which always contains a marked dip and conveys the complete information. In practical situations, system noise can be a common problem in MWP systems, as with any electronic system. The MWP systems are susceptible to noise that can arise from a range of sources, such as electronic components, amplifiers, fibers, and lasers. This noise can have a detrimental impact on sensing performance by introducing unwanted signals or degrading the quality of the MWP interrogation output. It is thus essential to mitigate the noise impact to ensure optimal sensing performance. Various techniques, including component selection, filtering, and signal processing algorithms, can be employed to reduce noise in MWP systems [7], [15], [16], [17]. However, implementing these techniques may lead to an increase in system complexity. To address this challenge, this work proposes the adoption of the DNN to directly learn from the raw interrogation output and perform estimation of the TGT.

B. Deep Learning Model

As a proof-of-concept of the proposed sensing scheme, we use the classic CNN to process the raw MWP interrogation outputs and develop a DL-assisted noise-resistant estimation model. CNN can be replaced with other neural networks as per specific practical applications. For example, recurrent neural networks can be employed when the sensors are deployed for monitoring a time-corrected process [18]. While CNNs are typically used for image recognition and computer vision tasks, their ability to automatically learn complex patterns without explicit feature engineering also makes them well-suited for processing one-dimensional time series data [19]. Fig. 3 illustrates the network structure of CNN and its training process. Each MWP interrogation process produces a LFM pulse that contains the resonance response and unwanted noise and interference, which is directly fed to the CNN, serving as one data point for deep learning processing. CNNs first use convolutional layers to extract features from the input data and pooling layers to reduce the dimensionality of the feature maps [20]. This allows the network to identify input patterns that are invariant to local variations, such as noise or small distortions [10], while further minimizing the influence of noise on the feature representations by summarizing local information across neighboring pixels or signal samples. Next, one or more fully connected layers learn to map those noise-invariant features extracted from the input data to the output neuron, which represents the estimated change in the TGT in this work. During supervised training, CNN computes the difference between its output prediction and the ground truth, resulting in a loss value to adjust the weights of the network through backpropagation. This process is repeated over many iterations until the loss is minimized and the network achieves satisfactory performance on a validation set. Additionally, the nonlinear activation functions enable CNNs to model complex nonlinear relationships, which can help remove ambiguity caused by highly correlated interference superimposed on the target resonance response.

The CNN model used in this work consists of four convolutional layers. The first convolutional layer has 64 filters of size 7×1, followed by two convolutional layers with 32 filters of size 5×1 and 3×1, respectively. The last convolutional layer uses 1 filter of size 1×1. Each of the first three convolutional layers incorporates a max-pooling layer with a kernel size of 2. The larger kernel sizes (7×1) are effective in capturing larger-scale patterns in the input spectrum and encourage the network to learn more generalized features, which can result in more efficient training, superior generalization performance, and less sensitivity to isolated noise. The two middle convolutional layers are used to extract detailed information and merge the lower-level features into higher-level representations. The final convolutional layer is used to reduce the number of channels as well as the network parameters, which helps to prevent overfitting. Following the convolutional layers, three fully connected layers with sizes of (2500, 512), (512, 256), and (256, 1) are used to facilitate regression. To further prevent overfitting, a dropout layer is added after the first fully connected layer. In practical applications, this one-dimension CNN model can be implemented on Field Programmable Gate Arrays with sub-nanosecond clock cycle times, enabling high-speed real-time data processing [21].

III. Experimental Section

The proposed DNN enhanced MWP sensing scheme is implemented as a proof-of-concept to measure the concentration of glycerol in distilled water in the presence of thermal interference and system noise, by using an on-chip air-cladded silicon-on-insulator (SOI) microdisk resonator. To record the temperature variation in the experiment, a reference microdisk coated with a thick layer of silicon dioxide on the same chip is utilized [22]. The proposed scheme can be employed for sensing other factors that can affect the effective index of the resonant mode, such as humidity, gases, biochemical molecules, and magnetic fields,

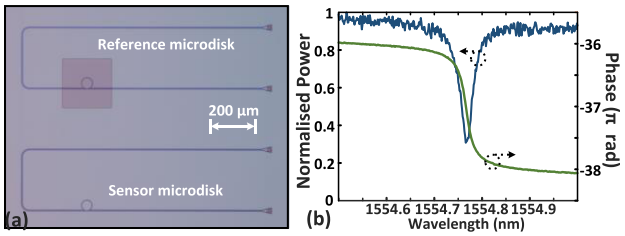

Fig. 4. (a) The SEM of the top view of the air-cladded microdisk used for the glycerol solution sensing and the SiO_2-coated microdisk used to monitor the thermal interference. (b) The optical power and phase transmission of the selected optical resonance of the air-cladded microdisk.

simply by adopting appropriate microresonator devices [1], [2], [3].

Fig. 4(a) shows the top view of the scanning electron microscopy of the microdisk resonators fabricated using the standard electron-beam lithography technology on the SOI wafer. The silicon devices are 220 nm thick and sit on top of a 2 μm thick buried oxide layer above a 725 μm thick silicon substrate. Both the sensor microdisk and the reference microdisk have a radius of 20 μm and are positioned 220 nm away from the bus waveguide, which has a width of 450 nm. By performing the plasma enhanced chemical vapor deposition, the reference microdisk is further coated with a layer of silicon dioxide that is nearly 700 nm thick to ensure that the resonant modes are not affected by glycerol concentration changes but are still sensitive to thermal interference. Vertical grating couplers (VGCs) are used for light coupling. Prior to measuring the glycerol solution, the temperature sensitivity of the reference microdisk was evaluated to be 82.66 pm/°C at approximately 1550 nm, using a Peltier unit and a temperature controller (Newport, model 325). Fig. 4(b) displays the optical power and phase transmissions of the selected optical resonant mode of the sensor microdisk under the air cladding condition, which were measured at a temperature of about 22 °C using an optical vector analyzer (Luna, OVA5000). The selected resonant mode is over-coupled and has a resonance wavelength of around 1554.77 nm. The measured ER and quality factor are approximately 5.1 dB and 43800, respectively.

The experimental setup of the glycerol concentration measurement against thermal interference and system noise is illustrated in Fig. 5(a). A tunable laser (Keysight, 81950A) was used as the laser source. The laser output was first directed to a polarization controller to ensure optimal transmission in the transverse electric (TE) mode. Before being injected into the microdisk sensor chip, the light was modulated in the DDMZM, which was configured with an electrical hybrid coupler and has a half-wave voltage of about 6.25 V, and was driven using a broadband (30 GHz) LFM pulse generated from an arbitrary waveform generator (AWG) (Keysight, M8195A). The DDMZM bias point was adjustable via a programmable power supply (Keysight, E3632A). A 90° fiber array with a gold coating was used to enable and maintain the light coupling between the fibers and the VGCs of the microdisk devices. The microdisk sensor chip, along with the glued fiber array, was placed in a homemade

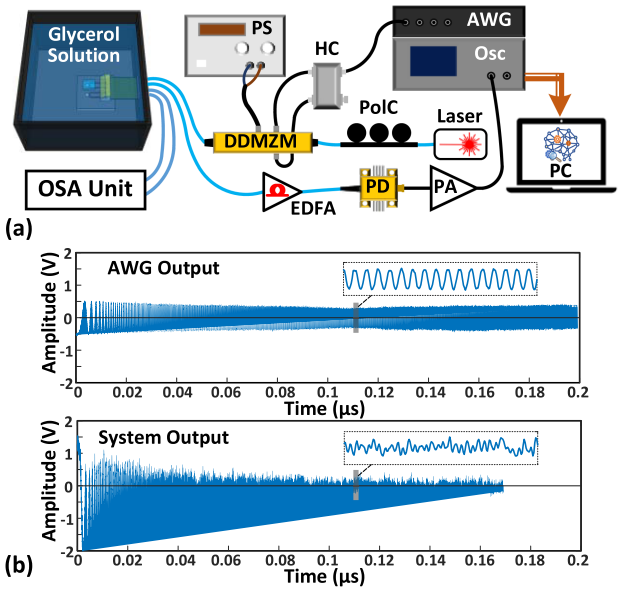

Fig. 5. (a) Experimental setup for the measurement of glycerol solution concentrations. (b) The experimental LFM pulse measured by the oscilloscope at the output of the AWG (top) and the PA (bottom) when the laser wavelength was set away from the resonance wavelength and the DDMZM was biased near the quadrature point. PS: Power supply; DDMZM: Dual-drive Mach-Zehnder modulator; HC: Hybrid coupler; AWG: Arbitrary waveform generator; Osc: Oscilloscope; PolC: Polarization controller; PA: Power amplifier; PD: Photodetector; EDFA: Erbium-doped fiber amplifier; PC: Personal computer; OSA: Optical spectrum analyzer.

solution container and left in the room condition without temperature control. To compensate for the optical coupling loss, an Erbium-doped fiber amplifier (EDFA) was inserted between the sensor microdisk output and a 40 GHz photodetector (Finisar). A high-speed oscilloscope (Tektronix, DPO77002SX), which has a sampling rate of 100 GS/s and a vertical resolution of 8 bits, was used to capture the transmitted LFM pulse signals, which were first amplified by an RF power amplifier after photodetection. The collected raw interrogation output data was processed on a computer using DNNs. To monitor the temperature variation of the chip immersed in the solution, the reference microdisk was connected to an optical spectrum analyzer unit.

The LFM pulse measured by the oscilloscope at the output of the AWG is shown at the top of Fig. 5(b). The experimental LFM pulse has a pulse width of 200 ns and consists of 20000 sampling points. The pulse repetition period was 400 ns, which hence allowed for an interrogation speed as fast as 2.5 MHz. Prior to the transmission through the MWP sensor system, the LFM pulse exhibited a clear sinusoidal waveform with an almost constant amplitude of approximately 0.5 V. The bottom portion of the figure displays the transmitted LFM pulse, which was measured at the MWP interrogation output while the DDMZM was biased close to the quadrature point and the laser wavelength was positioned away from the selected resonant mode. Due to the noise and nonlinearity introduced by the optical and electrical power amplifiers, as well as the imperfect response of the system components in practice, the transmitted LFM pulse shows a noisy envelope profile with a downward baseline and heavily distorted waveform shapes.

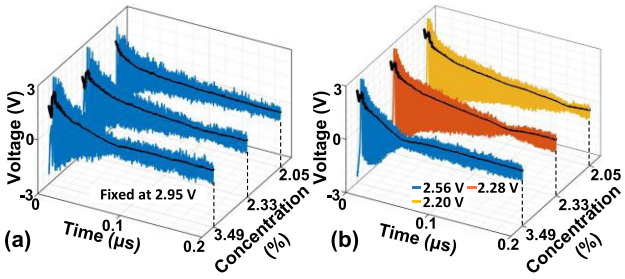

Fig. 6. Raw interrogation output against the upper root-mean-square envelope profiles (in black) collected at the glycerol concentration of 3.49%, 2.33%, and 2.05%, when the (a) fixed and (b) optimized DDMZM bias voltage was used, respectively.

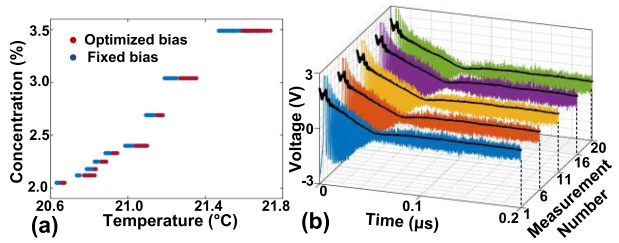

Fig. 7. (a) The chip temperature at each MWP interrogation using the fixed (blue dots) and optimized (red dots) DDMZM bias voltage, respectively; (b) The five out of twenty transmitted LFM pulse waveforms (the upper mean-square-error envelope profiles are in black) measured at the same glycerol concentration of 3.49%.

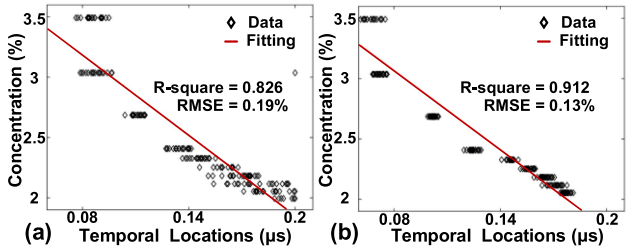

Fig. 8. The extracted envelope dip locations and the linear fitting results of the interrogation output at (a) fixed and (b) optimized DDMZM bias conditions.

IV. Results and Discussions

During the experiment, the sensor chip within the container was initially submerged in a glycerol solution with a concentration of 3.49%, while the laser wavelength was fixed at 1554.66 nm, which was on the longer wavelength side of the selected resonant mode. The glycerol solution was then successively diluted with distilled water to achieve concentrations of 3.04%, 2.69%, 2.41%, 2.33%, 2.25%, 2.18%, 2.12%, and 2.05%. At each of these nine different glycerol concentrations, 20 MWP interrogations were performed using the optimized bias voltages, and another 20 MWP interrogations were conducted with the DDMZM bias voltage fixed near the quadrature point. This results in a total of 180 raw transmitted LFM pulse waveforms collected for each case. At each MWP interrogation, the optical spectrum of the reference microdisk was simultaneously measured to determine the concurrent chip temperature.

First, the two groups of interrogation output were compared and analyzed using conventional methods. Fig. 6 shows the transmitted LFM pulse waveforms collected at three different glycerol concentrations using the fixed and optimized DDMZM bias voltages. To determine the temporal location of the resonance-dependent dip in the noisy pulse waveforms, the upper RMS envelope of each pulse waveform was extracted and processed with a moving average filter using a window size of 400, which is plotted in black against the original waveforms. Without the bias control, the envelope dip is wide and shallow at the initial concentration of 3.49%. However, it becomes increasingly challenging and eventually impossible to recognize the dip at the concentrations of 2.33% and 2.05%, respectively. In comparison, by using the optimized DDMZM bias voltage, an evident envelope dip was created and maintained throughout the measurement of different concentrations.

To illustrate the thermal interference during the experiment, the calculated chip temperature at each MWP interrogation is shown in Fig. 7(a). The chip temperature generally shows a continuous downward variation trend from 21.7 °C to 20.6 °C. This can be attributed to the addition of distilled water for dilution, while the temperature interference is arbitrary or subjected to the actual deployment environment of the sensor in practical applications. Five transmitted LFM pulse waveforms collected at the staring glycerol concentration by using the optimized bias voltages are presented in Fig. 7(b). From the first to the last MWP interrogation, the temporal location of the envelope dip shows an evident shift to larger time values, which indicates the effect of thermal interference on the interrogation output.

As a reference for comparison, by selecting the lowest point in the RMS envelope profile, the envelop dip locations for both the interrogations using fixed and optimized DDMZM bias voltages were extracted and processed with the conventional linear fitting. The results are shown in Fig. 8. Due to the presence of thermal interference and system noise, the envelope dip locations extracted under both bias conditions are not constant at each fixed glycerol concentration. The dip locations obtained from the MWP interrogation without bias control exhibit a more significant fluctuation, particularly at low experimental glycerol concentrations. In contrast, the change in the dip locations obtained under the optimized bias condition is more consistent with the thermal interference shown in Fig. 7(a). Without the bias control, the RMSE of the linear fitting model of the envelope dip locations is 0.19%, which is around 1.5-fold larger than that based on the interrogation using optimized bias voltages. These results highlight the advantage of the proposed time-domain MWP interrogation technique that uses optimized bias voltages. By creating a more distinct dip, it enables a higher interrogation resolution and accuracy against noise and interference.

Next, the DNN-assisted MWP sensing was implemented by training and testing a CNN using the raw interrogation output data obtained from MWP interrogation that utilizes optimized bias voltages. The experimental CNN uses the same structure shown in Fig. 3. The Adam optimizer [23] was employed to adjust the learning rate during the training. The LeakyReLU was

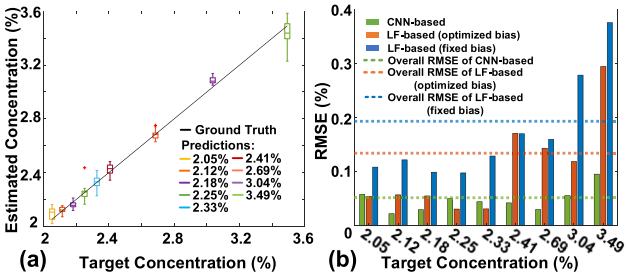

Fig. 9. (a) Boxplots of the estimated concentrations during the 5-fold cross validation; (b) The RMSE comparison of prediction models based on CNN and linear fitting (LF) of interrogation outputs using optimized and fixed bias conditions.

used as the activation function to address the vanishing gradient problem [24]. A batch size of 24 was selected to smooth the learning curve and improve the stability of the results. The 180 LFM pulse data were divided into training and testing datasets, with 80% used for training and 20% for testing. The noisy LFM pulse waveform, which consists of 20000 sample points, was used directly as the input of the CNN and compressed into a set of 2500 high-level features after passing through the four convolutional layers. 5-fold cross-validation was conducted to assess the model performance on different subsets of the dataset and ensure that it was not overfitting to a particular subset [13]. Fig. 9(a) shows the boxplots of the estimation results of the established CNN-based model during the 5-fold cross validation. The box in the middle represents the interquartile range of the distribution of the estimated values for each experimental glycerol concentration, which is the distance between the first (25^{th} percentile) and third (75^{th} percentile) quartiles. The line inside the box represents the median value. Compared with the distribution of the prediction results of the linear model shown in Fig. 8, except for a few outliers, the distribution of the predicted values of the CNN-based model for each experimental concentration is significantly more concentrated around the ground truth. Furthermore, the RMSE performance of the CNN-based model is compared with the linear models based on the envelope dip locations as handcrafted features in Fig. 9(b). The linear model using fixed bias conditions has a significantly larger RMSE at nearly all test concentrations. The linear model using optimized bias conditions shows a significantly improved accuracy but still has a high RMSE at the first four highest test concentrations. In comparison, the CNN-based model demonstrates a consistently low RMSE at each concentration. The CNN-based model achieved an overall RMSE of 0.05%, surpassing the performance of linear models based on handcrafted features by 2.6 and 3.8 times under optimized and fixed bias conditions, respectively. Furthermore, this achieved RMSE value is a 78-fold improvement compared to the utilization of metamaterial-based terahertz spectroscopy [25] and outperforms a commercially available Abbe refractometer-based glycerin concentration sensor (Misco, PA2022X) by a factor of 2. These results demonstrate the effectiveness of the proposed scheme and highlight the great potential of DNN in enhancing the MWP sensor performance. Additionally, for comparison purposes, the estimation model was built and tested through the same training, testing, and cross-validation processes by using a CNN with the same structure but a consistent filter size of 7×1 in each of its four convolutional layers, and a four-layer fully connected neural network with layer sizes of (2500, 512), (512, 256), and (256, 1). The achieved RMSEs for these two estimation models were approximately 0.07% and 0.17%, respectively. These findings further underscore the advantage of CNN in extracting informative features and highlight the importance of small kernel sizes in convolution to effectively detect and utilize intricate local features.

V. Conclusion

In summary, we have proposed a new MWP sensor which operates in the time domain and can achieve fast and accurate sensing in the presence of unwanted noise and interference, by adopting the DNN in processing the raw interrogation output without adding unnecessary system complexity. By using the LFM pulse and optimized modulator bias voltages in the optical sideband processing-based interrogation, the MWP sensor can achieve high speed and resolution. A detailed theoretical system model has been presented. As a proof-of-concept, the proposed scheme has been experimentally demonstrated for measuring the glycerol concentration in the presence of thermal interference and system noise by using a CNN to build the prediction model. The CNN-based model achieves a constant high estimation accuracy at all the experimental concentrations, which is 14 times better than the traditional estimation model established by linear fitting of the interrogation output without bias control. Those results prove the effectiveness of the proposed scheme and demonstrate its potential for high-end sensing applications.

Acknowledgment

The integrated devices were fabricated at Research Prototype Foundry at the University of Sydney. The authors would like to thank Mitchell Austin from the University of Sydney for his help with glycerol solution.

References

[1] X. Jiang, A. J. Qavi, S. H. Huang, and L. Yang, "Whispering-gallery sensors," *Matter*, vol. 3, no. 2, pp. 371–392, 2020.
[2] Y. Wang, S. Zeng, G. Humbert, and H. P. Ho, "Microfluidic whispering gallery mode optical sensors for biological applications," *Laser Photon. Rev.*, vol. 14, no. 12, 2020, Art. no. 2000135.
[3] Y. Zhi, X. C. Yu, Q. Gong, L. Yang, and Y. F. Xiao, "Single nanoparticle detection using optical microcavities," *Adv. Mater.*, vol. 29, no. 12, 2017, Art. no. 1604920.
[4] X. Tian et al., "High-resolution optical microresonator-based sensor enabled by microwave photonic sidebands processing," *J. Lightw. Technol.*, vol. 38, no. 19, pp. 5440–5449, Oct. 2020.
[5] X. Tian, L. Li, S. X. Chew, L. Nguyen, and X. Yi, "Cascaded optical microring resonator based auto-correction assisted high resolution microwave photonic sensor," *J. Lightw. Technol.*, vol. 39, no. 24, pp. 7646–7655, Dec. 2021.
[6] S. X. Chew et al., "Optoelectronic oscillator based sensor using an on-chip sensing probe," *IEEE Photon. J.*, vol. 9, no. 2, Apr. 2017, Art. no. 5500809.
[7] H. Deng, W. Zhang, and J. Yao, "High-speed and high-resolution interrogation of a silicon microdisk sensor based on microwave photonic filtering," *J. Lightw. Technol.*, vol. 36, no. 19, pp. 4243–4249, Oct. 2018.

[8] X. Tian, L. Li, L. Nguyen, R. A. Minasian, and X. Yi, "Microwave photonic sensor based on optical sideband processing with linear frequency-modulated pulse," in *Proc. IEEE Int. Topical Meeting Microw. Photon.*, 2022, pp. 1–4.

[9] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[10] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[11] S. Mohan, Z. Kadkhodaie, E. P. Simoncelli, and C. Fernandez-Granda, "Robust and interpretable blind image denoising via bias-free convolutional neural networks," in *Proc. Int. Conf. Learn. Representations*, 2020, pp. 1–22.

[12] D. Hepsiba and J. Justin, "Enhancement of single channel speech quality and intelligibility in multiple noise conditions using wiener filter and deep CNN," *Soft Comput.*, vol. 26, pp. 13037–13047, 2021.

[13] B. Xie and Q. Zhang, "Deep filtering with DNN, CNN and RNN," 2021, *arXiv:2112.12616v2*.

[14] W. Bogaerts et al., "Silicon microring resonators," *Laser Photon. Rev.*, vol. 6, no. 1, pp. 47–73, 2012.

[15] T. E. Darcie, A. Moye, P. F. Driessen, J. D. Bull, H. Kato, and N. A. F. Jaeger, "Noise reduction in class-AB microwave-photonic links," in *Proc. IEEE Int. Topical Meeting Microw. Photon.*, 2005, pp. 329–332.

[16] O. Daulay et al., "Ultrahigh dynamic range and low noise figure programmable integrated microwave photonic filter," *Nature Commun.*, vol. 13, no. 1, 2022, Art. no. 7798.

[17] P. Li, Z. Dai, L. Yan, and J. Yao, "A microwave photonic link with quadrupled capacity based on coherent detection and digital phase noise cancellation," *J. Lightw. Technol.*, vol. 40, no. 20, pp. 6845–6851, Oct. 2022.

[18] Y. Zhuang et al., "A hybrid data-driven and mechanistic model soft sensor for estimating CO2 concentrations for a carbon capture pilot plant," *Comput. Ind.*, vol. 143, 2022, Art. no. 103747.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[20] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," 2015, *arXiv:1511.08458v2*.

[21] R. Wu, X. Guo, J. Du, and J. Li, "Accelerating neural network inference on FPGA-based platforms—a survey," *Electronics*, vol. 10, no. 9, 2021, Art. no. 1025.

[22] M. K. Park et al., "Label-free aptamer sensor based on silicon microring resonators," *Sensors Actuators B Chem.*, vol. 176, pp. 552–559, 2013.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations*, 2015, pp. 1–15.

[24] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, "Yolov4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.

[25] W. Liang, J. Zuo, Q. Zhou, and C. Zhang, "Quantitative determination of glycerol concentration in aqueous glycerol solutions by metamaterial-based terahertz spectroscopy," *Spectrochimica Acta Part A: Mol. Biomol. Spectrosc.*, vol. 270, 2022, Art. no. 120812.