Optical performance monitoring in transparent fiber-optic networks using neural networks and asynchronous amplitude histograms

Jinsheng Xu[a], Jian Zhao[a, *], Sheng Li[a] and Tianhua Xu[a,b, **]

[a] Key Laboratory of Opto-Electronic Information Technology, Ministry of Education, School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China

[b]University of Warwick, Coventry CV4 7AL, United Kingdom

Email address: *enzhaojian@tju.edu.cn and **tianhua.xu@ieee.org

ABSTRACT:
Recently, technologies such as artificial neural networks (ANNs) and asynchronous amplitude histograms (AAHs) are widely employed in the optical performance monitoring (OPM). In this paper, the number of layers and the optimization of the neural networks are investigated to complete a series of monitoring tasks in optical communication systems with a large range of chromatic dispersion (CD) and optical signal-to-noise ratio (OSNR). It is shown that the monitoring accuracy has been significantly improved with such implementations. Numerical simulations, which have been conducted for six different modulation formats, demonstrate a good monitoring performance, with an average OSNR monitoring error of 0.1064 dB (99.49% monitoring accuracy) for the OSNR range of 10-40 dB and an average CD monitoring error of 3.3324 ps/nm (99.45% monitoring accuracy) for the CD range of 170-1870 ps/nm, respectively. Furthermore, the CD monitoring and the modulation format identification (MFI) have also been investigated for the system with a large range of dispersion. An average CD error of 16.6832 ps/nm (99.45% monitoring accuracy) and a 100% MFI accuracy have been achieved for all six modulation formats. This scheme is also verified experimentally with OSNR errors of 0.41 dB and CD errors of 15.21 ps/nm, respectively. The identification accuracy of the three modulation formats in the experimental system is also 100%.
Keywords: optical performance monitoring (OPM); optical signal-to-noise ratio (OSNR); chromatic dispersion (CD); neural network; asynchronous amplitude histogram

1. Introduction

With the rapid development towards high capacity, transparency and flexibility, future fiber-optic networks pose stronger requirements for the network reliability, and at the same time increase the complexity of the network management [1]. Optical performance monitoring (OPM) achieves the estimation and the acquisition of different physical parameters from transmitted signals and various components of optical networks. OPM is indispensable in ensuring the robust operation of optical networks and plays a key role in enhancing the flexibility and the efficiency of overall networks [2-4]. Conventional network performance monitoring and management in the electrical domain can no longer meet the requirements of next-generation dynamic heterogeneous optical networks [5-8]. In recent years, artificial neural networks (ANNs) have been applied for monitoring channel impairments and it has been shown that ANN is a powerful tool for the OPM [9-18]. Among these reports, the technique based on asynchronous amplitude histograms (AAHs) and ANNs, proposed by T. S. R. Shen and F. N. Khan, behaved simple and cost-effective [9,10]. Compared to other methods, this approach, employing AAHs as the NN input, can be applied to different modulation formats in coherent- and direct-detection communication systems, and does not require any manual feature extraction. However, in these reported works, the ranges of monitored parameters were quite small, and the monitoring accuracy had to be further improved. In this work, the simultaneous monitoring of chromatic dispersion (CD) and Optical signal-to-noise ratio (OSNR) with a larger monitoring range is achieved through multi-layer NNs, and the range of CD that can be monitored by this method is also studied. By changing the number of layers, optimizer, activation function and the number of bins, the monitoring error of neural network can be significantly reduced. The final calculated CD and OSNR monitoring errors using the above network

structure are much smaller than those reported in the literature, and the network also has a larger parameter monitoring range. Furthermore, the influence of the number of NN layers on the training time is studied using Google Colab. It is shown that the optimised multi-layer NN structure provides a higher monitoring accuracy while does not cost longer training time. In the end, experimental tests with two modulation formats have been carried out to verify the feasibility of the proposed method. This paper is organized as follows. In Section 2, numerical simulations of simultaneous monitoring on OSNR and CD have been performed in a transparent fiber-optic network with six different modulation formats, including return-to-zero on-off keying (RZ-OOK), non-return-to-zero (NRZ) OOK, chirped return-to-zero (CRZ) OOK, differential phase-shift keying (DPSK), differential quadrature phase-shift keying (DQPSK) and 16-ary quadrature amplitude modulation (16QAM), for the OSNR range of 10-40 dB and the CD range of 170-1870 ps/nm, respectively. Our results show that the average OSNR and CD monitoring errors are 0.1064 dB (99.49% monitoring accuracy) and 3.3324 ps/nm (99.45% monitoring accuracy), respectively, in the 10 Gbps fiber-optic network without the information of the modulation format. In Section 3, numerical simulations of CD monitoring have been carried out for the same six modulation formats with a larger CD range of 1870-5270 ps/nm, where an average CD monitoring error of 16.6832 ps/nm (99.45% monitoring accuracy) is achieved. In Section 4, numerical simulations of the modulation format identification (MFI) have been implemented, again for the same six modulation formats, and an estimation accuracy of 100% has been achieved. In Section 5, experimental demonstrations of MFI and simultaneous monitoring of CD and OSNR have been carried out. The conclusions of the paper are drawn in Section 6.

2. Numerical simulations of simultaneous monitoring on OSNR and CD

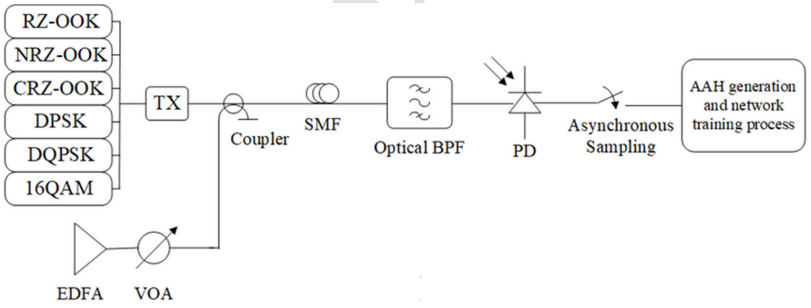


Fig. 1. Transmission setup for simultaneous monitoring of OSNR and CD. TX: transmitter; EDFA: erbium-doped fiber amplifier; VOA: variable optical attenuator; SMF: single-mode fiber; BPF: band-pass filter; PD: photodetector.

To demonstrate the performance of NN (trained with AAHs) in the simultaneous monitoring of CD and OSNR, numerical simulations have been performed using VPItransmissionMaker software, with a system setup shown in Fig. 1. Six modulation formats, including RZ-OOK, NRZ-OOK, CRZ-OOK, DPSK, DQPSK and 16QAM, are generated and transmitted over a SMF, all at a data rate of 10 Gbps. The EDFA is employed to introduce the amplified spontaneous emission (ASE) noise to the signal and the noise power is controlled by a VOA to adjust the OSNR in the range of 10 - 40 dB with a step size of 0.3 dB. The dispersion coefficient of the SMF is set as 17 ps/(nm▪km) and the total CD is in the range of 170 - 1870 ps/nm with a resolution of 17 ps/nm, controlled via the length of fiber. After the optical BPF, the signals are detected using a PD and are then asynchronously sampled at the receiver. The received signals, which are used to generate AAHs, contain 8448 amplitude samples. The abscissa of AAHs is the index of the bin, and the ordinate is the number of corresponding occurrences. Different

from the existing methods of manually extracting features, AAHs contain information about the amount of different impairments in the system, and AAHs are sensitive to changes in OSNR and CD of the transmission link. Therefore, the AAHs represented by a vector of number of occurrences can be used as the input neurons and the outputs of the NN are the actual OSNR and CD values of the system. The NN can automatically extract the features in the AAHs, and continuously optimize the parameters of the network during the training process, so that the two outputs of the network gradually approach the actual CD and OSNR values. The data set includes six different modulation formats, and each has 10201 combinations of CD and OSNR values. Within the 61206 sets of data, we randomly select 40000 components as the training data, 10000 components as the validation data and 10000 components as the test data, for the training and the application of NN.

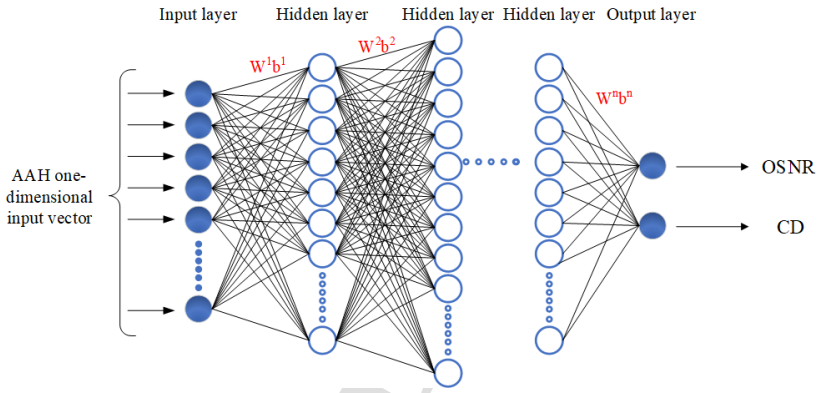


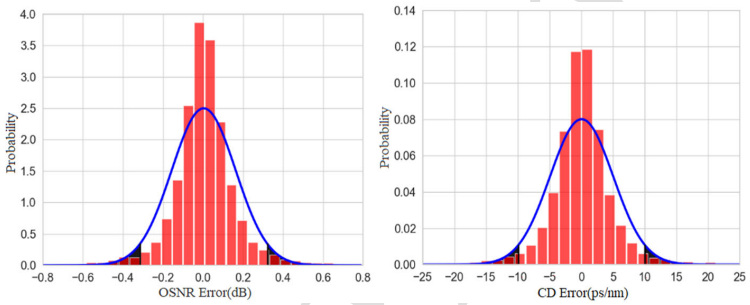
Fig. 2. Structure of the NN used for simultaneous monitoring of OSNR and CD.

The NN used in this work is a fully connected scheme with multiple hidden layers, as shown in Fig. 2. The input of the NN is the number of amplitude samples that fall into each bin interval. The number of bins is chosen as 80, meaning that the input layer of the NN has 80 neurons. The training data are fed in batches, with a batch size of 32. The activation function of the NN is the Rectified Linear Unit (ReLU), since it requires less complexity in training and optimization and can show a better performance [19,20]. The training data set [X_{Train}, Y_{Train}] is employed to optimize the NN parameters (w and b) by minimizing the mean-square-error (MSE) $|Y_{Train}-Y_{Predict}|^2$ over the whole training set. The optimizer used in the NN is the adaptive moment estimation (Adam). The Adam algorithm integrates the ideas of Gradient Descent, Momentum, Adagrad, RMSprop and their improved variants [21]. The advantages of the Adam algorithm include low computing power, small memory requirements, insensitivity to gradient scaling and easy adjustment of hyperparameters. The learning rate of Adam is initially set as 0.001. When the loss starts to behave stable, the learning rate will drop to one tenth of its initial value. NNs with different numbers of hidden layers are applied to perform the simultaneous monitoring of OSNR and CD, and results are shown in Table 1.

Table 1 Results of simultaneous monitoring of OSNR and CD with different numbers of hidden layers

The numbers of hidden layers	2	4	6	8	10
Average OSNR error(dB)	0.2922	0.1427	0.1257	0.1064	0.1163
Average CD error(ps/nm)	10.2799	4.7763	4.0070	3.3320	3.3871
Training time(s)	427.1315	583.6901	747.2634	989.0715	1231.8016

The average errors of the OSNR and the CD in Table 1 are calculated based on $\sum_1^{10000} |Y_{Test} - Y_{Prediction}| /10000$. It is found that the NN shows the best performance when the number of hidden layers is 8. The training time is calculated using Google Colab with a graphics processing unit (GPU) Tesla T4. Although it can be seen from Table 1 that the more hidden layers of the NN the longer the training time, the 8-hidden layer NN takes the least time when the NNs have the same loss. This arises from the fast convergence speed and fewer required numbers of epochs in the training of the 8-hidden layer NN. Since the training process is implemented offline, the training time is not the major consideration in defining the number of NN layers. When the number of hidden layers is 8, the CD and the OSNR monitoring accuracy, calculated as $(1 - \frac{\sum_1^{10000}|(Y_{Test} - Y_{Prediction})/Y_{Test}|}{10000}) \times 100\%$, are 99.45% and 99.49%, respectively.



(a) (b)

Fig. 3. (a) Histogram and normal distribution fitting of the OSNR estimation error using the 8-hidden layer NN. (b) Histogram and normal distribution fitting of the CD estimation error using the 8-hidden layer NN.

Figure 3 depicts the histograms and the correponding normal distribution fittings of the CD and the OSNR estimation errors using the 8-hidden layer NN, respectively. It can be seen that most of the OSNR errors lie between -0.18 and 0.18 dB and more than 95% errors are situated between -0.32 and 0.32 dB. Similarly, most of the CD errors lie between -5 and 5 ps/nm, and more than 95% CD errors are located between -10 and 10 ps/nm.

3. CD monitoring for a wider range of dispersion

To investigate the performance of the NN (trained with AAHs), the transmission system in Fig. 1 is employed to perform the CD monitoring with a larger range of dispersion. The OSNR is in the range of 10-40 dB with a step size of 1.0 dB, and the CD value is in the range of 170-8670 ps/nm with a step size of 34 ps/nm. The data set again includes six modulation formats, and each scheme has 7781 combinations of CD and OSNR values. We divide the CD values, between 170 ps/nm and 8670 ps/nm, into five groups to verify the CD monitoring behavior in each group. From each group of data, we randomly select 5000 components as the training set, 2000 components as the validation set, and 2000 components as the test set for the training and the application of NN. The training set is also imported in batches, with a batch size of 16. The number of the hidden layers is again selected as 8. The number of neurons in the output layer is 1, since only the CD monitoring is performed.

Table 2 Average CD monitoring error in each group

Group index	1	2	3	4	5
Range of CD (ps/nm)	170-1870	1870-3570	3570-5270	5270-6970	6970-8670
Average CD error (ps/nm)	2.8253	7.5979	39.0034	423.6310	426.8650

The average CD monitoring error in each group is listed in Table 2. Figure 4 describes the scatter plots of predicted CD versus actual CD for group 1-5. It can be seen in Fig. 4 that the CD monitoring in the first two groups works well, while the CD monitoring in the third group shows some degradation. The CD monitoring cannot be performed in the last two groups (group 4 and 5). This is because when the CD is over large, AAHs are not sensitive to the variation of CD any more.

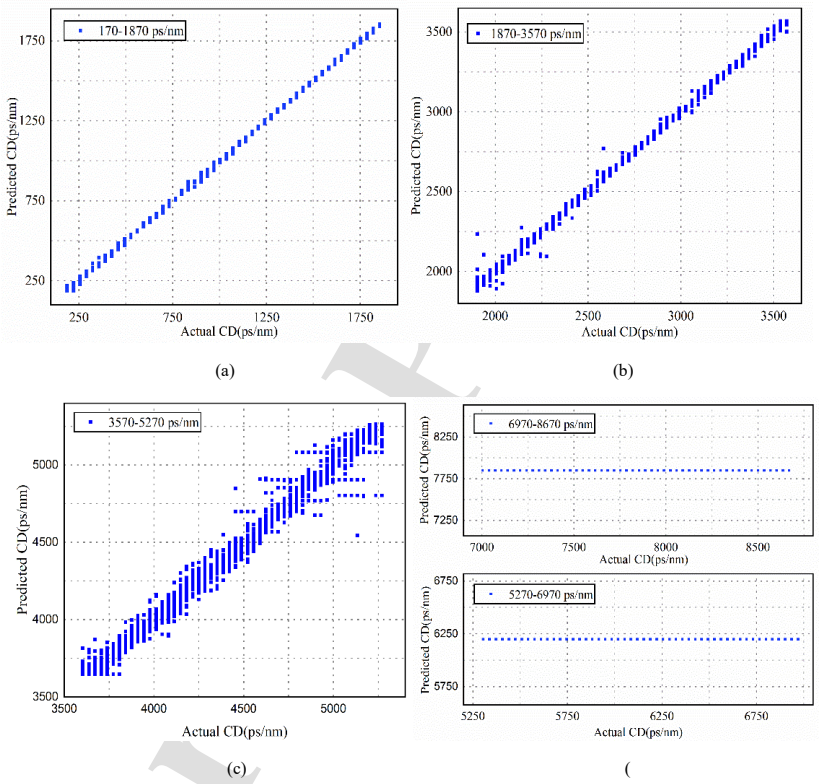


Fig. 4. (a)-(c) Scatter plots of predicted CD versus actual CD for group 1-3, respectively; (d) Scatter plots of predicted CD versus actual CD for group 4 and 5.

Since it is already known that the CD monitoring using NNs (trained with AAHs) is only applicable to the dispersion of up to 5270 ps/nm, the impact of the number of bins on the CD monitoring performance is then investigated for the CD range of 1870-5270 ps/nm.

Table 3 Average CD errors under different number of bins

The number of bins	80	100	120	140	160	180	200
Average CD error (ps/nm)	29.1243	24.2481	22.8859	20.2782	18.4844	17.5490	16.6832

Table 3 shows the average CD monitoring errors when the numbers of bins are 80, 100, 120, 140, 160, 180 and 200, respectively. When the number of bins increases from 80 to 100, the average CD monitoring error shows a significant reduction, however, when the number of bins continues rising, the average CD monitoring error decreases more and more slowly. In general, the input features become more obvious as the number of bins increases. This is more conducive to the NN training. Therefore, a larger number of bins can reduce the average monitoring error of CD. However, it will require more input neurons, and will result in the increased computational complexity. Since there is a trade-off between the CD monitoring error and the computational complexity, the number of bins should not be set too much. When the number of bins is 200, the CD monitoring accuracy is 99.47%.

4. MFI for a large range of dispersion

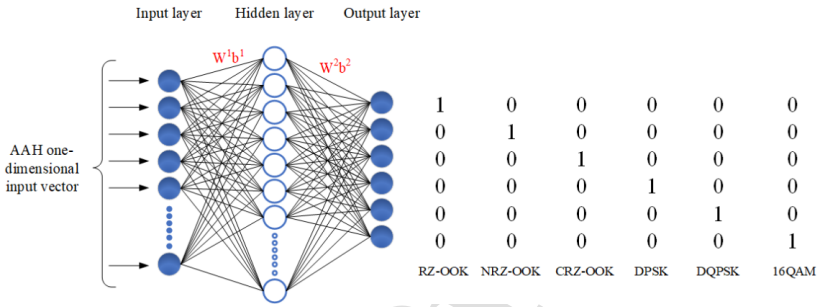
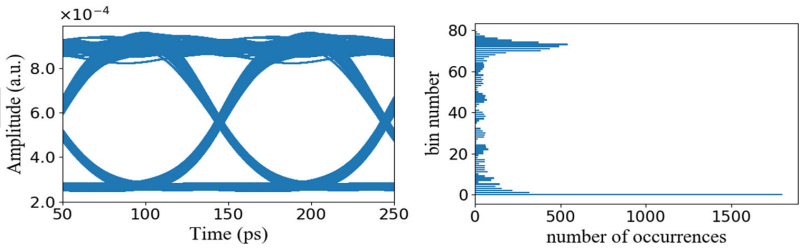


Fig. 5. Structure of the NN used for MFI.

The NN used for MFI employs a fully connected structure, as shown in Fig. 5, where only a single hidden layer is applied. The numbers of neurons in the input and the output layers are determined by the number of bins (e.g. 80 in the test here) in AAHs and the number of considered modulation formats (e.g. 6 in this paper), respectively. The activation function used for the first two layers is ReLU, and the activation function used for the last layer is Softmax. The output of Softmax function characterizes the relative probability between different categories and this activation function has wide application senarios in multi-classification tasks [22]. The output of the NN is a one-dimensional vector with six probabilities, and the classification result is determined by the index of the highest probability. The loss function in Softmax is the cross entropy which can measure the degree of difference in random variables with different probability distributions. The smaller the cross entropy value is, the better the model prediction behaves. The optimizer used in the NN is Adam. The data set includes six different modulation formats, each scheme has 4681 combinations of CD (in the range of 170-5270 ps/nm with a resolution of 34 ps/nm) and OSNR (in the range of 10-40 dB with a step size of 1.0 dB). Similar as previous operations, within the 28086 sets of data, we randomly select 16000 components as the training set, 6000 components as the validation set, and 6000 components as the test set for the training and the application of NN.



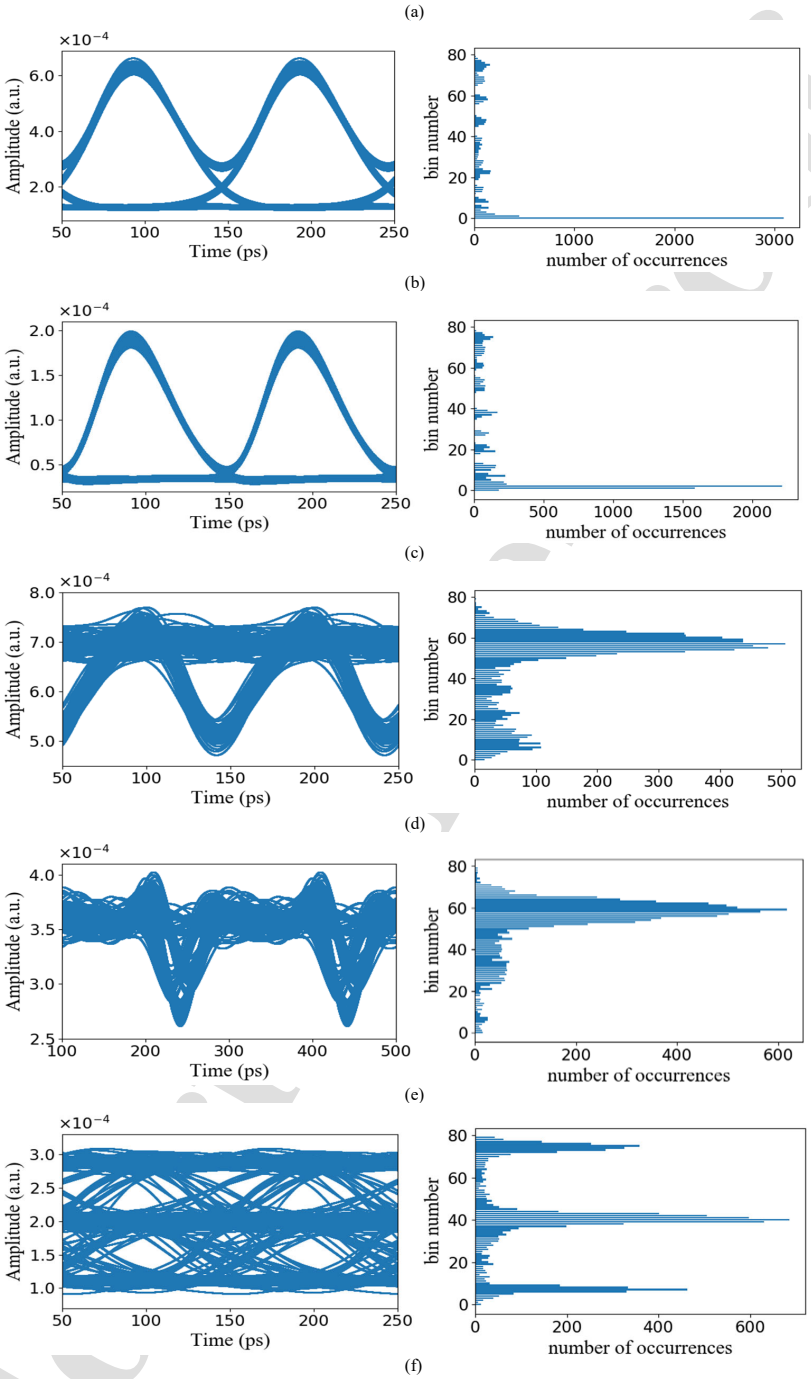


Fig. 6. Eye-diagrams and corresponding AAHs for (a) NRZ-OOK; (b) RZ-OOK; (c) CRZ-OOK; (d) DPSK; (e) DQPSK; (f) 16QAM. The CD is 170 ps/nm and the OSNR is 30 dB.

Figure 6 shows eye-diagrams and corresponding AAHs for different modulation formats. AAHs for these six modulation formats behave significantly different, and thus such AAHs can be used as input

features of the NN. The overall results for MFI using the NN proposed in Fig. 5 are summarized in Table 4. It is clear from the table that all the six modulation formats have been well classified with an overall accuracy of 100 % despite a considerable range of OSNR and CD.

Table 4 Indentification accuracies of six modulation formats

Actual Modulation Format	Indentified Modulation Format					
	RZ-OOK	NRZ-OOK	CRZ-OOK	DPSK	DQPSK	16QAM
RZ-OOK	100 %	-	-	-	-	-
NRZ-OOK	-	100 %	-	-	-	-
CRZ-OOK	-	-	100 %	-	-	-
DPSK	-	-	-	100 %	-	-
DQPSK	-	-	-	-	100 %	-
16QAM	-	-	-	-	-	100 %

Table 5 lists some comparisons between the poposed method and other machine learning (ML) techniques using specific features in other literature. Compared with the previous work, the monitoring range and accuracy of the method we proposed have been significantly improved [23].

Table 5 Comparisons between the poposed method and other machine learning techniques

Features domain	ML techniques	Computation Complexity	Modulation formats	OSNR monitoring range (dB)	Classification accuracy	Monitoring accuracy	Reference
CDF	SVM	SVM and DNN have comparable time complexity.	DP-QPSK	5-30	-	MAE=0.24 dB	[24]
			DP-16QAM			MAE=0.24 dB	
			DP-64QAM			MAE=0.24 dB	
Constellation images	CNN	CNN achieves higher time complexity.	QPSK	15-30	100 %	Accuracy>95%	[25]
			8PSK		100 %		
			8QAM		100 %		
			16QAM		100 %		
			32QAM		100 %		
			64QAM		100 %		
AAH	DNN	-	NRZ-OOK	10-40	100 %	MAE=0.11 dB Accuracy>99%	This work
			RZ-OOK		100 %		
			CRZ-OOK		100 %		
			DPSK		100 %		
			DQPSK		100 %		
			16QAM		100 %		

5. Experimental setup for OPM

Simultaneous monitoring of OSNR and CD in the experimental system was performed first. Fig.7 (a) and (b) show the experimental setup of the 10 Gbps NRZ-OOK and QPSK signals for the simultaneous

monitoring of CD and OSNR using NNs and AAHs. In Fig. 7 (a), a tunable laser centered at 1550 nm is externally modulated (with an intensity modulator, Fujitsu FTM7920FBA) by 10 Gb/s pseudo-random bit sequence (PRBS) signals with a pattern length of 2^{15}-1 to generate the NRZ-OOK signals. An EDFA and a VOA are employed to produce different levels of ASE noise to adjust the OSNR values. The noise output is combined with the modulated signals through a 3-dB coupler. The combined (noise and data) signals are amplified by another EDFA and are then launched into the SMF with different lengths, which is used to introduce different amount of CD. The signals are further tapped using a 10:90 coupler to implement a conventional OSNR measurement with an optical spectrum analyzer (OSA), while signals in the other branch (90%) pass through a bandpass filter (BPF) with a 3 dB bandwidth of 0.6 nm. Optical signals are detected using a 10 GHz photodiode. The PD converts the received optical signals into electrical signals, which are then asynchronously sampled by the digital signal analyzer (DSA, Tektronix DSA 8300). In principle, the electrical signal can be asynchronously sampled at a rate much slower than the symbol rate. The sampling rate of DSA we set in this experiment is 80 Gbps to reduce the sampling time. In Fig. 7(b), a tunable laser at 1550 nm is modulated (with the QPSK modulator, SHF 46213D) by a 5 Gbaud signal generator (Anritsu MG3694C) to produce the QPSK signals. The rest of the setup and devices are roughly the same as decribed in Fig. 7 (a). The range of the OSNR monitoring is from 15 dB to 30 dB with a resolution of 0.5 dB. The dispersion coefficient of the SMF in this experiment is about 15.93 ps/(nm·km) and the length of the SMF varies from 0 to 46.5 km with a step size of 3.1 km. In other words, the CD monitoring range is from 0 to 740.75 ps/nm with a resolution of 49.38 ps/nm. For each specific system condition and signal format, five samples including an input feature vector and two output values have been collected. The experimental data set contains two modulation formats, and each scheme has 2480 combinations of CD and OSNR values. Within the 4960 sets of data, we randomly select 3800 components as the training data, 500 components as the validation data and 500 components as the test data, for the training and the application of NN.

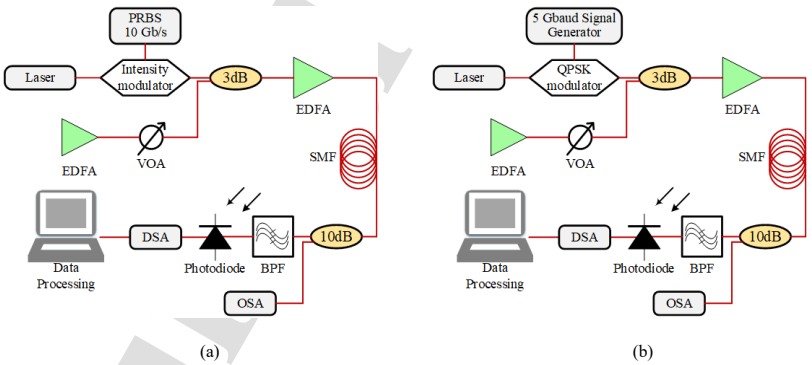


Fig. 7. Experimental setup of the (a) NRZ-OOK (b) QPSK transmission system.

Using a three-layer neural network, the simultaneous monitoring of the CD and the OSNR has been implemented experimentally. Figure 8 (a) and (b) show the experimental results of the OSNR monitoring. It can be seen that most of the OSNR errors lie between -0.6 and 0.6 dB and more than 95% errors are situated between -0.8 and 0.8 dB. The average OSNR monitoring error of the 500 test sets is 0.48 dB.

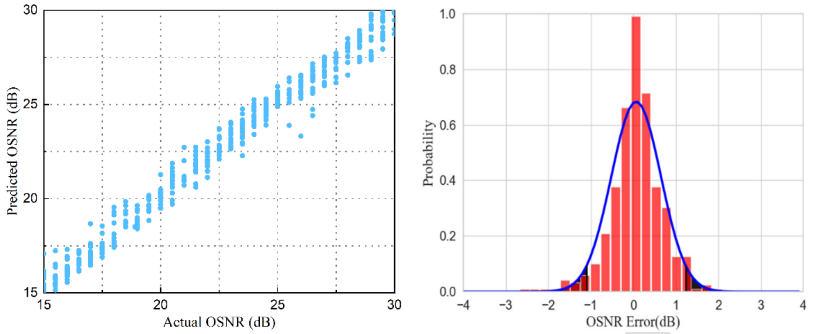


Fig. 8. (a) Scatter plot of predicted OSNR versus actual OSNR in the experiment.(b) Histogram and normal distribution fitting of the OSNR estimation error in the experiment.

Figure 9 (a) and (b) show the experimental results of the CD monitoring. Figure 9(a) describes the scatter plot of the predicted CD versus the actual CD. It can be seen from Fig. 9(b) that most of the CD errors lie between -20 and 20 ps/nm and more than 95% errors are situated between -27 and 27 ps/nm. The results indicate that the average CD monitoring error of the 500 test sets is 15.86 ps/nm. It is found that the estimation errors in the experiments are larger than those in the simulations. This is because that there are not sufficient amount of data collected in the experiments due to the limitations in the lab. Since the dispersion resolution in the experiment is 49.38 ps/nm, the data of many intermediate dispersion values cannot be collected. Using more than 4000 sets of experimental data, we have experimentally proved the feasibility of the proposed method.

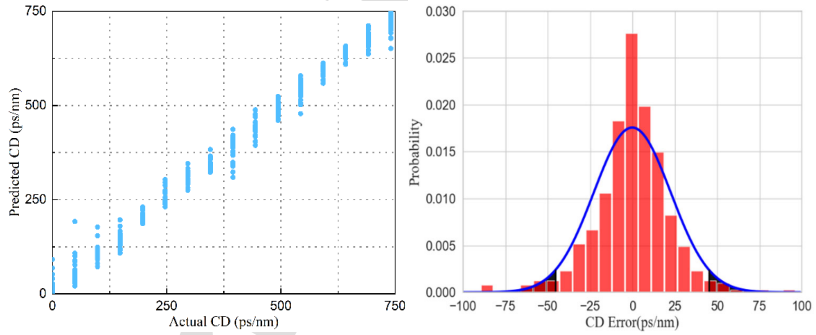


Fig. 9. (a) Scatter plot of predicted CD versus actual CD in the experiment.(b) Histogram and normal distribution fitting of the CD estimation error in the experiment.

After completing simultaneous monitoring of OSNR and CD, the task of MFI using the experimental data was carried out. The experimental data here includes the data of QPSK and NRZ-OOK mentioned above, as well as the data of binary phase shift keying (BPSK). We randomly select 4000 components as the training set, 500 components as the validation set, and 500 components as the test set for the training and the application of NN with 3 layers. The numbers of neurons in the input and the output layers are 80 and 3, respectively. The overall results for MFI of the experimental data using the NN are summarized in Table 6. It is clear from the table that all the three modulation formats have been well classified with an overall accuracy of 100 %. The experimental system and the simulation system have

the same identification results. The classificaition results obtained by experiment and simulation are the same.

Table 6 Indentification accuracies of three modulation formats in the experimental system

Actual Modulation Format	Indentified Modulation Format		
	NRZ-OOK	QPSK	BPSK
NRZ-OOK	100 %	-	-
QPSK	-	100 %	-
BPSK	-	-	100 %

6. Conclusion

In this paper, OPM tasks have been investigated based on the application of NNs, which are specifically trained with AAHs. This method is not only applicable to the 10 Gbps fiber-optic network, but also is applicable to higher-speed optical networks. When the number of hidden layers is 8, the simultaneous monitoring of CD and OSNR shows the best performance, with an average OSNR error of 0.1064 dB and an average CD error of 3.3324 ps/nm. It is also found that this proposed monitoring scheme, based on AAHs trained NNs, works well for the CD range of up to 5270 ps/nm, and the increase in the number of bins can reduce the average monitoring error. In the CD range of 1870-5270 ps/nm, the increase in the number of bins can make input features more distinct. This is beneficial to the training of the NNs. With the further increase of the CD (larger than 5270 ps/nm), the input features corresponding to each label become almost the same and cannot be employed for the NN training any more. Numerical simulations also indicate that the accuracy of the MFI is 100% for the CD range of 170-5270 nm/ps. Furthermore, the feasibility of the proposed method using NNs and AAHs has been experimentally verified. Therefore, OPM tasks in fiber-optic networks can be implemented in an accurate, effiencient and low-cost manner using the proposed NN-AAH structure.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Zhao, Y. Liu and T. Xu, “Advanced DSP for Coherent Optical Fiber Communication,” Applied Sciences, vol. 9, no. 19, 2019.

[2] Z. Dong, F. N. Khan, Q. Sui, K. Zhong, C. Lu and A. P. T. Lau, “Optical performance monitoring: A review of current and future technologies,” IEEE Journal of Lightwave Technology, vol. 34, no. 2, pp. 525-543, 2016.

[3] J. Zhao, Z. Li, D. Liu, L. Cheng, Chao Lu, and H. Y. Tam, “NRZ-DPSK and RZ-DPSK Signals Signed Chromatic Dispersion Monitoring Using Asynchronous Delay-Tap Sampling,” Journal of Lightwave Technology, vol. 27, no. 22, pp. 5295-5301, 2009.

[4] J. Zhao, A. P. T. Lau, K. K. Qureshi, Z. Li, C. Lu and H. Y. Tam, “Chromatic Dispersion Monitoring for DPSK Systems Using RF Power Spectrum,” Journal of Lightwave Technology, vol. 27, no. 24, pp. 5704-5709, 2009.

[5] Z. Li and G. Li, “Chromatic dispersion and polarization-mode dispersion monitoring for RZ-DPSK signals based on asynchronous amplitude-histogram evaluation,” Journal of Lightwave Technology, vol. 24, no. 7, pp. 2859-2866, 2006.

[6] Z. Li, C. Lu, Y. Wang and G. Li, “In-service signal quality monitoring and multi-impairment discrimination based on asynchronous amplitude histogram evaluation for NRZ-DPSK systems,” IEEE Photonics Technology Letters, vol. 17, no. 9, pp. 1998-2000, 2005.

[7] C. Yu and Y. Yu, “Optical performance monitoring in fiber transmission systems based on electrical sampling technique,” 16th International Conference on Transparent Optical Networks, pp. 1-4, 2014.

[8] Y. Yu, B. Zhang and C. Yu, “Optical signal to noise ratio monitoring using single channel sampling technique,” Optics Express, vol. 22, no. 6, pp. 6874-6880, 2014.

[9] T. S. R. Shen, K. Meng, A. P. T. Lau and Z. Y. Dong, “Optical performance monitoring using artificial neural network trained with asynchronous amplitude histograms,” IEEE Photonics Technology Letters, vol. 22, no. 22, pp. 1665-1667, 2010.

[10] F. N. Khan, Y. Zhou, A. P. T. Lau and C. Lu, “Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks,” Optics Express, vol. 20, no. 11, pp. 12422-12431, 2012.

[11] X. X. Wu, J. A. Jargon, R. A. Skoog, L. Paraschis and A. E. Willner, “Applications of artificial neural networks in optical performance monitoring”, Journal of Lightwave Technology, vol. 27, no. 16, pp. 3580-3589, 2009.

[12] V. Ribeiro, M. Lima, A. Teixeira, “Comparison of optical performance monitoring techniques using artificial neural networks”, Neural Computing and applications, vol. 23, no. 3, pp. 583–589, 2013.

[13] J. A. Jargon, X. X. Wu, H. Y. Choi, Y. C. Chung and A. E. Willner, “Optical performance monitoring of QPSK data channels by use of neural networks trained with parameters derived from asynchronous constellation diagrams”, Optics Express, vol. 18, no. 5, pp. 4931-4938, 2010.

[14] F. N. Khan, T. S. R. Shen, Y. Zhou, A. P. T. Lau, and C. Lu, “Optical performance monitoring using artificial neural networks trained with empirical moments of asynchronously sampled signal amplitudes,” IEEE Photonics Technology Letters, vol. 24, no. 12, pp. 982-984, 2012.

[15] Y. Huang, Y. Chen, and J. Yu, “Optical performance monitoring of 56Gbps optical PAM4 signal using artificial neural networks,” Asia Communications and Photonics Conference, 2017.

[16] J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones and D. Zibar, “Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals,” Journal of Lightwave Technology, vol. 35, no. 4, pp. 868-875, 2017.

[17] J. A. Jargon, X. Wu and A. E. Willner, “Optical performance monitoring using artificial neural networks trained with eye-diagram parameters,” IEEE Photonics Technology Letters, vol. 21, no. 1, pp. 54-56, 2009.

[18] T. S. R. Shen, Q. Sui and A. P. T. Lau, “OSNR monitoring for PM-QPSK systems with large inline chromatic dispersion using artificial neural network technique,” IEEE Photonics Technology Letters, vol. 24, no. 17, pp. 1564-1567, 2012.

[19] K. Jarrett, K. Kavukcuoglu, M. Ranzato and Y. LeCun, “What is the best multi-stage architecture for object recognition?,” IEEE 12th International Conference on Computer Vision, pp. 2146-2153, 2009.

[20] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, pp. 315–323 2011.

[21] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization", Proceedings of the International Conference of Learning Representation,” pp. 1-41, 2015.

[22] D. Ciregan, U. Meier and J. Schmidhuber, “Multi-column deep neural networks for image classification,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 3642-3649, 2012.

[23] W. S. Saif, M. A. Esmail, A. M. Ragheb, T. A. Alshawi and S. A. Alshebeili, “Machine Learning Techniques for Optical Performance Monitoring and Modulation Format Identification: A Survey,” IEEE Communications Surveys & Tutorials, vol. 22, no. 4, pp. 2839-2882, 2020.

[24] X. Lin, O. A. Dobre, T. M. N. Ngatched, Y. A. Eldemerdash and C. Li, “Joint Modulation Classification and OSNR Estimation Enabled by Support Vector Machine,” IEEE Photonics Technology Letters, vol. 30, no. 24, pp. 2127-2130, 2018.

[25] D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, “Intelligent constellation diagram analyzer using convolutional neural network-based deep learning,” Optics Express, vol. 25, no. 15, pp. 17150-17166, 2017.