Effect of bandwidth of direct detection receiver on multiparameter optical performance monitoring

Luo, Huaijian, Huang, Zhuili, Du, Xinwei, Yu, Changyuan

Huaijian Luo, Zhuili Huang, Xinwei Du, Changyuan Yu, "Effect of bandwidth of direct detection receiver on multiparameter optical performance monitoring," Proc. SPIE 11555, Real-time Photonic Measurements, Data Management, and Processing V, 115550H (10 October 2020); doi: 10.1117/12.2575168

SPIE.

Event: SPIE/COS Photonics Asia, 2020, Online Only

Invited Paper

Effect of bandwidth of direct detection receiver on multiparameter optical performance monitoring

Huaijian Luo[1], Zhuili Huang[2], Xinwei Du[1,3], and Changyuan Yu[1]

[1]Photonics Research Centre, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China
[2]Key Laboratory of Optoelectronic Technology & Systems (Ministry of Education), College of Optoelectronic Engineering, Chongqing University, Chongqing, 400044, China
[3]Division of Science and Technology, Beijing Normal University-Hong Kong Baptist Univerisity United International College, Zhuhai, 519087, China

ABSTRACT

Our proposed optical performance monitoring (OPM) technique, based on the multi-task learning, is able to realize modulation format identification (MFI), baud rate identification (BRI), chromatic dispersion identification (CDI), and optical signal-to-noise ratio (OSNR) estimation simultaneously. This OPM technique can be used in the intermediate nodes of optical networks, which is cost-effective since it can monitor systems by direct detection. To further reduce the cost, PD and ADC with low bandwidth are used. Therefore, we investigate the effect of different bandwidths of direct detection receiver on the performance of OPM, to find the optimal low bandwidth of receiver to achieve relatively high OPM performance. The comparison experiment between different bandwidths of receivers has been carried out where signals with two formats, quadrature phase shift keying (QPSK) and 16 quadrature amplitude modulation (16QAM), two baud rates, 14 GBaud and 28 GBaud, and three CD situations, 0 ps/nm, 858.5 ps/nm, and 1507.9 ps/nm, are adopted. And the test bandwidths of receiver are 0.5 GHz, 1 GHz, 2.5 GHz, 5 GHz, 7.5 GHz, 10 GHz, and the original 33 GHz. It is found that the model with 5 GHz bandwidth receiver has the relatively highest performance, except the model with original 33 GHz bandwidth receiver. The identification accuracies of model with 5 GHz receiver are 99.92%, 99.11%, and 99.94% for MFI, BRI, and CDI, respectively. The OSNR estimation error of this model is 0.594 dB.

Keywords: Optical performance monitoring, Bandwidth of receiver, Multi-parameter, Machine learning

1. INTRODUCTION

Nowadays, the optical fiber transport network becomes more and more complicated by applying schemes with high spectral efficiency, such as dense wavelength division multiplexing (DWDM) and elastic optical network (EON).[1] This brings a significant challenge to monitor the condition of signal and maintain the quality of service (QoS). To deal with this issue, optical performance monitoring (OPM) techniques were proposed to monitor signals through the network. The key parameters worth monitoring are modulation format (MF), baud rate (BR), chromatic dispersion (CD), and the most important one, optical signal-to-noise ratio (OSNR).

Machine learning (ML) has attracted a lot of attention these years. The deep learning (DL), as one of the ML branch, which has various applications in different research areas. Compared with conventional methods, the ML-assisted OPM techniques can build the mapping between the input and the output directly without formula derivation. This means the dedicated step of constructing model can be skipped. Many ML algorithms are being applied in the field of OPM, like principal component analysis[2] (PCA), support vector machine[3,4] (SVM), kernel-based methods[5], artificial neural network[6–9](ANN). Moreover, some DL algorithms have good performance as well. Such as the deep neural network[10] (DNN), convolutional neural network[11,12] (CNN), and long short-term memory[13] (LSTM). In addition to monitoring one parameter of the signal, DL models can

Further author information: (Send correspondence to Changyuan Yu)
Huaijian Luo: E-mail: huaijian.luo@connect.polyu.hk
Changyuan Yu: E-mail: changyuan.yu@polyu.hk

Real-time Photonic Measurements, Data Management, and Processing V, edited by Ming Li, Bahram Jalali, Mohammad Hossein Asghari, Proc. of SPIE Vol. 11555, 115550H · © 2020 SPIE
CCC code: 0277-786X/20/$21 · doi: 10.1117/12.2575168

monitor various parameters at the same time by combining with multi-task learning, like measuring MF and OSNR[14–18], CD and OSNR[19], MF, BR and OSNR[20].

In this paper, we not only proposed a ML-assisted OPM technique that can monitor multiple parameters simultaneously with direct detection but also investigated the effect of the bandwidth of the receiver on OPM performance. The aim of this proposed method is to use low-bandwidth receivers along the internodes of the communication network to achieve cost-effective monitoring for multiple parameters. The following paper is elaborated as follows. Section 2 introduces what the dataset consists of, and the detail of the neural network model; Section 3 shows the experimental setup and dataset and the corresponding results. Finally, the conclusions are shown in section 4.

2. METHOD

In this section, the generation processes of two types of samples are firstly introduced. Then, the structure and related information of multiple inputs and multi-task learning model is discussed, especially the advantages of multi-input and multi-task model.

2.1 AADTP and AAH

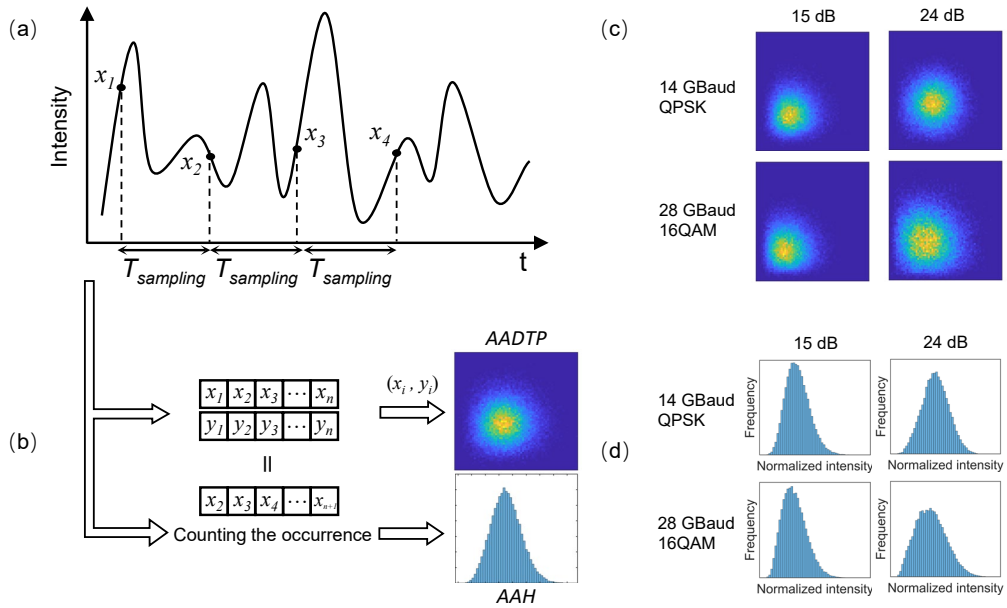

Figure 1. Schematic diagram of generation of AADTP and AAH process with (a) direct detection and (b) two algorithms. (c) AADTPs generated by signals with different formats, baud rates, and OSNR values. (d) AAHs generated by signals with different formats, baud rates, and OSNR values.

Adaptive asynchronous delay tap sampling plot (AADTP) and asynchronous amplitude histogram (AAH) are features we choose to represent the condition of signals. Both of them can be generated from the same waveform sampled by ADC. This generation process is shown in Fig. 1 (a) and (b).

For AADTP, it evolved from conventional ADTP[21,22], a 2D matrix. The reason why using AADTP is that there are some limits in the traditional ADTP. First is the symbol rate is needed when generating ADTP. For example, the delay times applied in ADTP are usually 1/2 or 1/4 symbol intervals, so we have to adjust the delay line according to the symbol rate. However, the information of signals is always unknown. The second problem is that we cannot adjust the delay line dynamically when signals with different baud rates pass through the fiber. Therefore, we develop the algorithm of generating AADTP by fixing the delay as the sampling interval. By this way, the delay time is unified among all signals, which is only dependent on the sampling rate of ADC.

The generation process is as follows: (1) Sampling: get a sequence x_i ($i \in [1, 50000]$). (2) Setting delay: set $y_i = x_{i+j} (j \in N^*)$. Here we set $j = 1$. (3) Composing coordinates: compose two sequences as (x_i, y_i). (4) Counting: count the distribution of coordinates as AADTP. We can see that the styles of AADTP are different with distinct formats, symbol rates and OSNRs from Fig. 1 (c). Thus, features of AADTP can be used to recognize the parameters of signals.

AAH can be treated as a 1D vector because the x-axis denotes the discrete normalized intensity, which can be seen as the index of vector, and the y-axis is the number of occurrences for each intensity interval, which represent the value of the vector. The generation process is relatively easier compared to AADTP because only two steps are required. The first step is the same, sampling the sequence x_i. Second, counting the occurrence each the amplitude interval that x_i belongs to. Similarly, AAH has its own features on the distribution, as illustrated in Fig. 1 (d). In our work, we combine the AADTP and AAH to increase the information, which can further develop the performance of the neural network model compared to using single AADTP or AAH.

2.2 Multiple inputs and multi-task learning

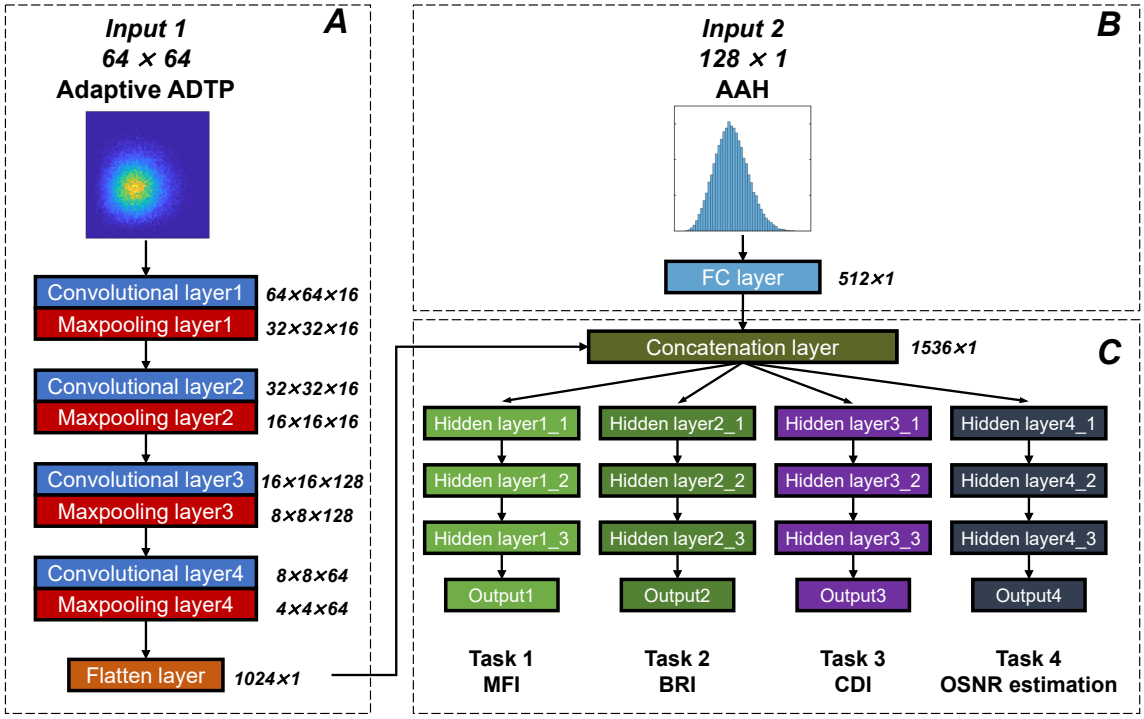

Figure 2. The structure of multiple inputs and multi-task learning model.

After getting the dataset, the next step is to construct the multiple inputs and multi-task learning model. After designing and optimizing the hyperparameters of the neural network, the structure can be expressed by Fig. 2. This neural network is made of three parts. Part A mainly consists of convolutional layers and max-pooling layers and its function is extracting features from the AADTP. The structure of part B is relatively simple for it only has one fully connected layer to extract information from the AAH. Part C concatenates the output of A and B into one layer and shares neurons for four separate tasks. Task 1-3 are classifications, and their loss function is cross-entropy, shown in Eq. (1),

$$L_k = -\frac{1}{M} \sum_{i=1}^{M} \sum_{c=1}^{N} y_{i,c} \log \left(\hat{y}_{i,c} \right) \tag{1}$$

$$L_k = -\frac{1}{M}\sum_{i=1}^{M}\left(y_i \log\left(\hat{y}_i\right) + \left(1-y_i\right)\log\left(1-\hat{y}_i\right)\right) \tag{2}$$

where the subscript k represents the number of tasks, M denotes the total number of samples for this task and N is the number of types for classification. If $N = 2$, then the Eq. (1) will be degenerated to Eq. (2). Task 4 is the regression, whose loss function is mean absolute error (MAE) shown in Eq. (3)

$$L_{\text{mae}} = \frac{1}{M}\sum_{i=1}^{M}\left|y_i - \hat{y}_i\right| \tag{3}$$

$$L = \lambda_1 L_1 + \lambda_2 L_2 + \lambda_3 L_3 + \lambda_4 L_{\text{mae}} \tag{4}$$

Although root mean square error (RMSE) is used more frequently, the MAE exceeds RMSE as the metric to evaluate the regression error of model[23]. And Eq. (4) describe the combination of the total loss function. It should be noted that the weight factor λ decide the performance of the neural network model, which will be elaborated in the section 3.2. The advantage of multi-task learning is these four tasks can share common hyperparameters in the concatenated layer[24]. Thus, this can enhance the accuracies of each task.

3. EXPERIMENTS

In this paper, the multi-task learning model is trained by datasets generated by waveforms sampled from the experimental setup below. The first subsection introduces the experimental process. Then, the second subsection tells details about the dataset. In the final part, results about the effect of the bandwidth of the receiver on the performance of the neural network are discussed.

3.1 Experimental setup

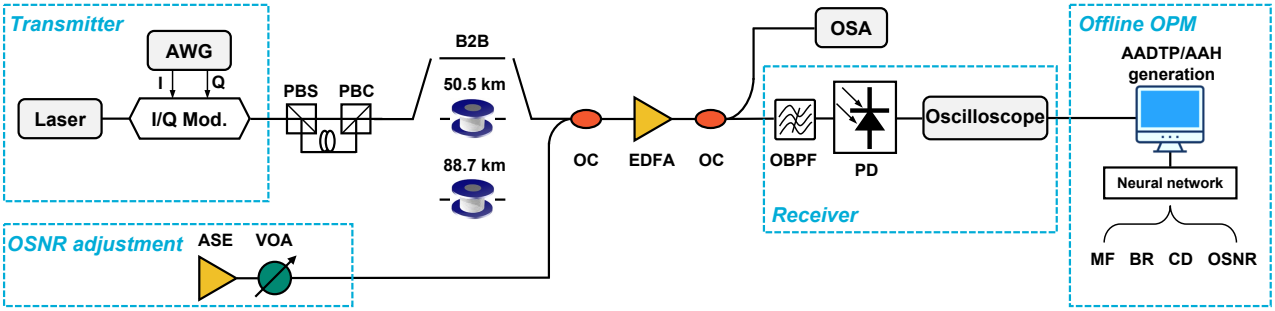

Figure 3. Schematic diagram of the experimental setup.(AWG: arbitrary waveform generator, ASE: amplified spontaneous emission noise, VOA: variable optical attenuator, PBS: polarization beam splitter, PBC: polarization beam coupler, OC: optical coupler, EDFA: erbium-doped fiber amplifiers, OBPF: optical bandpass filter, PD: photodetector, OSA: optical spectral analyzer)

In Fig. 3, the experimental setup mainly consists of five parts. Except for the transmission part, the other four parts in blue dash boxes are transmitter part, OSNR adjustment part, receiver part, and offline OPM part, respectively. In the transmitter part, we can adjust the AWG to change the modulation format and symbol rate. For OSNR adjustment part, the power of ASE noise can be attenuated by the VOA to tune the OSNR. The PBS and PBC in the transmission part are used to create the pseudo dual polarization. Different length of fibers can introduce distinct CD. The offline OPM part trains the model based on the dataset generated by waveforms from the receiver part, and OSNR labels come from the OSA. It should be noted that the bandwidth of PD is 40 GHz and the sampling rate of oscilloscope is 80 GSa/s.

3.2 Datasets and training details

In the experiment, we applied two modulation formats (QPSK and 16QAM), two symbol rates (14 GBaud and 28 GBaud), three kinds of CD (0 ps/nm, 858.5 ps/nm, and 1507.9 ps/nm). For QPSK format, the range of OSNR is from 10 to 24 dB, and the granularity is 1 dB. At the same time, the range is 15-29 dB for 16QAM with same granularity. After the process by algorithms in MATLAB, the sampled waveforms are transformed into AADTPs and AAHs. The number of AADTP and AAH per dB is 200, which means that we have 36000 samples of AADTPs and AAHs. We divide the whole dataset into three sets: training set, validation set and test set. The proportions of them are 80%, 10%, and 10%, separately. The validation set used in training can clearly show if the model is overfitting. And the test set is used to present the overall performance of the model.

The multi-task learning model is based on the Keras, a python deep learning library, which uses Tensorflow as the backend. We set "same" as the padding way of convolutional layers to keep the size of the inputted matrix unchanged. The activation function of every layer is the exponential linear unit (ELU) instead of the rectified linear unit (RELU) because of ELU s faster convergence speed and higher generalization capacity. To get rid of overfitting, we apply the L2 regularization with coefficient equals to 0.001. Another critical point is how to select the values of λ in the loss function. Because it denotes the weight of each task, and it affects the overall performance of prediction. The loss of OSNR regression is usually the largest. Thus we set λ_4 to 1 and select $\lambda_1 = 0.189$, $\lambda_2 = 0.158$, $\lambda_3 = 0.188$ by using another algorithm, which is called Bayesian optimization[25].

To validate the effect of the bandwidth of receiver on the OPM performance, we apply digital filters (10 GHz, 7.5 GHz, 5 GHz, 2.5 GHz, 1 GHz, and 0.5 GHz) on the original waveform to simulate the filtering effect. The reason why using digital filters is because of the lack of low-bandwidth receivers. After the filtering, we can get 7 types of datasets including the original one.

3.3 Results

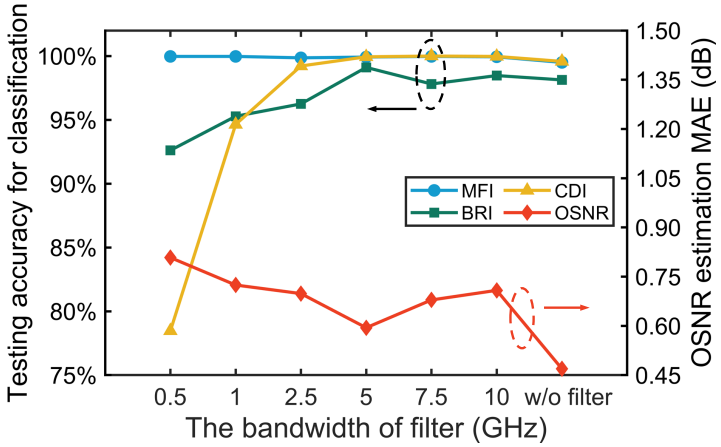

Figure 4. Performance of models on datasets with different bandwidth filtering

It should be noted that all the neural network models are trained for ten times to eliminate the initial randomness. For this best model, the accuracies of MFI, BRI, CDI are 99.50%, 98.13%, 99.58%, separately. The OSNR estimation MAE is 0.4691 dB. Then, the training process is concentrated on the dataset with filtering. Fig. 4 shows the accuracies of the MFI, BRI, CDI and the MAE values of OSNR estimation of different models trained from filtered datasets. It can be seen that the MFI is the most stable task because accuracies over different filtered dataset are almost 100%. For BRI and CDI, accuracies increase as the raise of bandwidth and accuracies of BRI can be maintained over 90%. The MAE values over the filtered datasets are higher than that without filtering. From the aspect of performance, we can find the performance over 5 GHz filtered dataset is the best (MFI: 99.92%, BRI: 99.11%, CDI: 99.94%, OSNR MAE: 0.5944 dB) and even close to the performance over the dataset without filtering. In summary, low-bandwidth receivers can be used in the internodes of optical networks, and relatively good OPM performance can still be achieved meanwhile.

4. CONCLUSIONS

In this paper, we study the effect of the bandwidth of direct detection receiver on multiparameter optical performance monitoring with the help of multi-task learning model. In the experiment, signals with two types of modulation formats (QPSK, 16QAM), two baud rates (14 GBaud, 28 GBaud), three kinds of chromatic dispersions (0 ps/nm, 858.5 ps/nm, and 1507.9 ps/nm) and several OSNR values are tested. After training the model with datasets with $0.5 \sim 10$ GHz filtering, we find the performances of some models from low-bandwidth schemes are still good. Among them, the 5 GHz model is the best one, whose accuracies of MFI, BRI and CDI are 99.92%, 99.11%, and 99.94%, respectively. And the total MAE of OSNR is 0.5944 dB. This performance is very close to that of the model trained by dataset without filtering. The real low-bandwidth receiver should be tested in the future to validate this conclusion further. Meanwhile, the bandwidth of PD and sampling rate of ADC should be investigated separately to find a combination that is cost-effective and of good performance.

ACKNOWLEDGMENTS

This work was supported by Grant 15211619 from HK RGC GRF and 1-ZE5K from HK PolyU, and 62005030 from NSFC Program.

REFERENCES

[1] Shen, G., Guo, H., and Bose, S. K., "Survivable elastic optical networks: survey and perspective," *Photonic Network Communications* **31**(1), 71–87 (2016).

[2] Tan, M. C., Khan, F. N., Al-Arashi, W. H., Zhou, Y., and Lau, A. P. T., "Simultaneous optical performance monitoring and modulation format/bit-rate identification using principal component analysis," *Journal of Optical Communications and Networking* **6**(5), 441–448 (2014).

[3] Wang, D., Zhang, M., Li, Z., Cui, Y., Liu, J., Yang, Y., and Wang, H., "Nonlinear decision boundary created by a machine learning-based classifier to mitigate nonlinear phase noise," in [*2015 European Conference on Optical Communication (ECOC)*], 1–3, IEEE (2015).

[4] Wang, D., Zhang, M., Cai, Z., Cui, Y., Li, Z., Han, H., Fu, M., and Luo, B., "Combatting nonlinear phase noise in coherent optical systems with an optimized decision processor based on machine learning," *Optics Communications* **369**, 199–208 (2016).

[5] Anderson, T. B., Kowalczyk, A., Clarke, K., Dods, S. D., Hewitt, D., and Li, J. C., "Multi impairment monitoring for optical networks," *Journal of Lightwave Technology* **27**(16), 3729–3736 (2009).

[6] Khan, F. N., Shen, T. S. R., Zhou, Y., Lau, A. P. T., and Lu, C., "Optical performance monitoring using artificial neural networks trained with empirical moments of asynchronously sampled signal amplitudes," *IEEE Photonics Technology Letters* **24**(12), 982–984 (2012).

[7] Kashi, A., Zhuge, Q., Cartledge, J., Borowiec, A., Charlton, D., Laperle, C., and O'Sullivan, M., "Fiber nonlinear noise-to-signal ratio monitoring using artificial neural networks," in [*2017 European Conference on Optical Communication (ECOC)*], 1–3, IEEE (2017).

[8] Caballero, F. V., Ives, D., Laperle, C., Charlton, D., Zhuge, Q., OSullivan, M., and Savory, S. J., "Machine learning based linear and nonlinear noise estimation," *Journal of Optical Communications and Networking* **10**(10), D42–D51 (2018).

[9] Kashi, A. S., Zhuge, Q., Cartledge, J. C., Etemad, S. A., Borowiec, A., Charlton, D. W., Laperle, C., and O'Sullivan, M., "Nonlinear signal-to-noise ratio estimation in coherent optical fiber transmission systems using artificial neural networks," *Journal of Lightwave Technology* **36**(23), 5424–5431 (2018).

[10] Li, J., Wang, D., and Zhang, M., "Low-complexity adaptive chromatic dispersion estimation scheme using machine learning for coherent long-reach passive optical networks," *IEEE Photonics Journal* **11**(5), 1–11 (2019).

[11] Zhang, W., Zhu, D., He, Z., Zhang, N., Zhang, X., Zhang, H., and Li, Y., "Identifying modulation formats through 2d stokes planes with deep neural networks," *Optics express* **26**(18), 23507–23517 (2018).

[12] Wang, D., Zhang, M., Li, J., Li, Z., Li, J., Song, C., and Chen, X., "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Optics express* **25**(15), 17150–17166 (2017).

[13] Wang, Z., Yang, A., Guo, P., and He, P., "Osnr and nonlinear noise power estimation for optical fiber communication systems using lstm based deep learning technique," *Optics express* **26**(16), 21346–21357 (2018).

[14] Yi, A., Yan, L., Liu, H., Jiang, L., Pan, Y., Luo, B., and Pan, W., "Modulation format identification and osnr monitoring using density distributions in stokes axes for digital coherent receivers," *Optics express* **27**(4), 4471–4479 (2019).

[15] Xiang, Q., Yang, Y., Zhang, Q., and Yao, Y., "Joint and accurate osnr estimation and modulation format identification scheme using the feature-based ann," *IEEE Photonics Journal* **11**(4), 1–11 (2019).

[16] Wan, Z., Yu, Z., Shu, L., Zhao, Y., Zhang, H., and Xu, K., "Intelligent optical performance monitor using multi-task learning based artificial neural network," *Optics express* **27**(8), 11281–11291 (2019).

[17] Wang, D., Wang, M., Zhang, M., Zhang, Z., Yang, H., Li, J., Li, J., and Chen, X., "Cost-effective and data size–adaptive opm at intermediated node using convolutional neural network-based image processor," *Optics express* **27**(7), 9403–9419 (2019).

[18] Khan, F. N., Zhong, K., Zhou, X., Al-Arashi, W. H., Yu, C., Lu, C., and Lau, A. P. T., "Joint osnr monitoring and modulation format identification in digital coherent receivers using deep neural networks," *Optics express* **25**(15), 17767–17776 (2017).

[19] Wang, C., Fu, S., Wu, H., Luo, M., Li, X., Tang, M., and Liu, D., "Joint osnr and cd monitoring in digital coherent receiver using long short-term memory neural network," *Optics express* **27**(5), 6936–6945 (2019).

[20] Cheng, Y., Fu, S., Tang, M., and Liu, D., "Multi-task deep neural network (mt-dnn) enabled optical performance monitoring from directly detected pdm-qam signals," *Optics express* **27**(13), 19062–19074 (2019).

[21] Anderson, T., Dods, S., Kowalczyk, A., Clarke, K., Hewitt, D., and Li, J., "Optical performance monitoring based on asynchronous delay-tap sampling," in [*Optical Performance Monitoring: Advanced Techniques for Next-Generation Photonic Networks*], 175–192, Elsevier (2010).

[22] Yu, Y., Zhang, B., and Yu, C., "Optical signal to noise ratio monitoring using single channel sampling technique," *Optics Express* **22**(6), 6874–6880 (2014).

[23] Willmott, C. J. and Matsuura, K., "Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance," *Climate research* **30**(1), 79–82 (2005).

[24] Caruana, R., "Multitask learning," *Machine learning* **28**(1), 41–75 (1997).

[25] Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and De Freitas, N., "Taking the human out of the loop: A review of bayesian optimization," *Proceedings of the IEEE* **104**(1), 148–175 (2015).