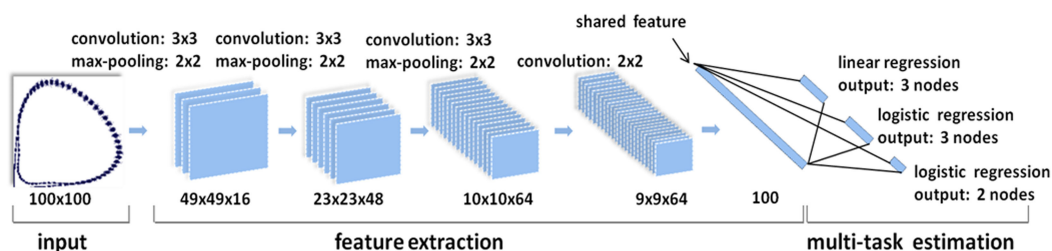


Joint Optical Performance Monitoring and Modulation Format/Bit-Rate Identification by CNN-Based Multi-Task Learning

Volume 10, Number 5, September 2018

Xiaojie Fan
Yulai Xie
Fang Ren
Yiying Zhang
Xiaoshan Huang
Wei Chen
Tianwen Zhangsun
Jianping Wang






A novel technique is proposed for joint multi-impairment optical performance monitoring (OPM) with bit-rate and modulation format identification (BR-MFI) in next-generation heterogeneous optic communication networks by convolution neural network (CNN)-based deep multi-task learning (MTL) on asynchronous delay-tap sampling (ADTS) phase portraits. Instead of treating the monitoring and identification tasks as separate problems, a novel MTL technique is used to joint optimization of them utilizing the ability of feature extraction and feature sharing.

DOI: 10.1109/JPHOT.2018.2869972

1943-0655 © 2018 IEEE

Joint Optical Performance Monitoring and Modulation Format/Bit-Rate Identification by CNN-Based Multi-Task Learning

Xiaojie Fan ¹, Yulai Xie,² Fang Ren ¹, Yiyang Zhang,¹
Xiaoshan Huang,¹ Wei Chen,¹ Tianwen Zhangsun,¹
and Jianping Wang ¹

¹School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

²Hitachi (China) Research and Development Co., Ltd. Beijing 100083, China

DOI:10.1109/JPHOT.2018.2869972

1943-0655 © 2018 IEEE. Translations and content mining are permitted for academic research only.

Personal use is also permitted, but republication/redistribution requires IEEE permission.

See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Manuscript received July 28, 2018; revised September 5, 2018; accepted September 9, 2018. Date of publication September 13, 2018; date of current version October 5, 2018. This work was supported in part by the Natural National Science Foundation of China (NSFC) under Grant 61605004, in part by the Fundamental Research Funds for the Central Universities, China No. FRF-TP-16-046A1, and in part by the State Key Laboratory of Advanced Optical Communication Systems Networks, China. Corresponding author: Fang Ren (e-mail: renfang@ustb.edu.cn).

Abstract: A novel technique is proposed for joint multi-impairment optical performance monitoring (OPM) with bit-rate and modulation format identification (BR-MFI) in next-generation heterogeneous optic communication networks by convolution neural network (CNN)-based deep multi-task learning (MTL) on asynchronous delay-tap sampling phase portraits. Instead of treating the monitoring and identification tasks as separate problems, a novel MTL technique is used to joint optimization of them utilizing the ability of feature extraction and feature sharing. Compared with principal component analysis-based pattern recognition algorithm, CNN-based MTL achieves the better accuracies and has a shorter processing time (~56 ms). The combination signals of three modulation formats and two bit rates under various impairments are used in numerical simulation. For OPM, the results show monitoring of optical signal-to-noise ratio, chromatic dispersion, and differential group delay with root-mean-square error of 0.73 dB, 1.34 ps/nm, and 0.47 ps, respectively. Similarly, for BR-MFI, even in the case of limited training data, 100% accuracies can be achieved. Additionally, the effects of training data size, task weights, and model structure on CNN-based MTL performance are comprehensively studied. The proposed technique can intelligently analyze the signals of future heterogeneous optic communication networks, and the analysis results are helpful for better management of optical networks.

Index Terms: Optical performance monitoring (OPM), bit-rate and modulation format identification (BR-MFI), convolution neural network (CNN), deep multi-task learning (MTL), asynchronous delay-tap sampling (ADTS).

1. Introduction

Future optical networks are designed to be dynamic and heterogeneous. In order to meet the various demands of end-users, The optical networks need the ability of transmitting various signals with different modulation formats and bit rates. Due to the complexity of the heterogeneous optic networks, optical signals may traverse different paths and accumulate a variety of transmission

impairments [1]. Hence, it is necessary to deploy enough OPM devices at the intermediate nodes of dynamic optical fiber communication network for real-time monitoring of channel impairments.

Similarly, BR-MFI can provide additional information to select an appropriate carrier recovery module or improve OPM precision. In recent years, the combination of OPM and BR-MFI has become a trend of development. Various machine learning algorithms (MLAs) have been applied for optical communications, including back-propagation artificial neural network (BP-ANN) [2], [3], k-nearest neighbors (KNN) [4], as well as support vector machine (SVM) [5]. For joint OPM and BR-MFI, a PCA-based pattern recognition on asynchronous delay-tap plot (ADTP) method was proposed [6]. Likewise, the same method is proposed for joint OSNR monitoring and MFI [7]. The first step of PCA-based pattern recognition algorithm is to extract fixed size feature vectors of the given image and all images in the reference dataset. Then, a feature vector is selected from the reference dataset, which has the minimum Euclidean distance with the feature vector of the given image. Finally, we estimate the result of the given image using the labels (impairment values, bit-rate and modulation format) corresponding to the selected feature vectors from the reference dataset.

Unfortunately, The MLAs were lack of the ability of feature extraction and sharing. Specifically, the raw form of the data cannot be processed by a MLA model, meanwhile the design of a feature extractor needs rich domain experience. Furthermore, the problem of joint OPM and BR-MFI is comprised of impairments monitoring task (which is a regression task), bit rate identification task and format identification task (which are classification tasks). We believe that the performance of the impairments monitoring task can be affected by some related factors. For instance, low bit-rate signals are less affected than high bit-rate signals under the same impairments. Effectively discovering and utilizing the information provided by the related factor such as bit-rate would help in monitoring signal impairments more accurately. Also, different modulation formats have different responses to the same impairments. The information provided by the modulation format can be used to improve the precision of impairments monitoring. However, MLAs cannot share useful feature information among different tasks. In order to avoid the drawbacks of MLAs, the more advanced algorithms are needed to not only automatically extract features but also share features among different tasks.

In recent years, more and more attention has been paid to deep learning, among which the MTL is popular in many applications [8]–[12]. Our work introduces the concept of MTL to jointly optimize impairments monitoring with BR-MFI. The aim of the MTL is to improve the performance of related tasks by extracting and sharing useful information among them [13]. As far as we know, this is the first attempt to use MTL for joint optimization of OPM and BR-MFI tasks, though MTL is not new in the field of computer vision. Few works combine optical communication systems with deep learning. One closest work to us is Ref. [14], which doing MFI and OSNR monitoring by CNN. However, their work suffers from the following drawbacks. (i) They only provide OSNR information as the quality evaluation of the signal, regardless of other important factors such as CD and DGD. (ii) Their work is limited to format identification and without bit-rate identification. (iii) They treat OSNR monitoring as a classification problem, but the impairment values should be continuous variables. And the CNN model is limited in the ability of feature sharing.

In this paper, a novel technique is proposed for joint OPM and BR-MFI in heterogeneous optical networks by employing CNN-based MTL on phase portraits generated using ADTS. Recently, due to the potential to be applied to impairments monitoring and BR-MFI, ADTS has attracted wide attention. The phase portraits can be processed by the CNN-based MTL, which has the ability of feature extraction and feature sharing, for the purpose of joint optimization of OPM and BR-MFI without knowing other prior information. In order to verify the effectiveness of the proposed technique, numerical simulations are performed for 10/20 Gbps non-return-to-zero on-off keying (NRZ-OOK), 10/20 Gbps return-to-zero on-off keying (RZ-OOK), 10/20 Gbps non-return-to-zero differential phase shift keying (NRZ-DPSK) signals. The joint OPM and BR-MFI show good performance under the help of CNN-based MTL's ability of feature extraction and feature sharing.

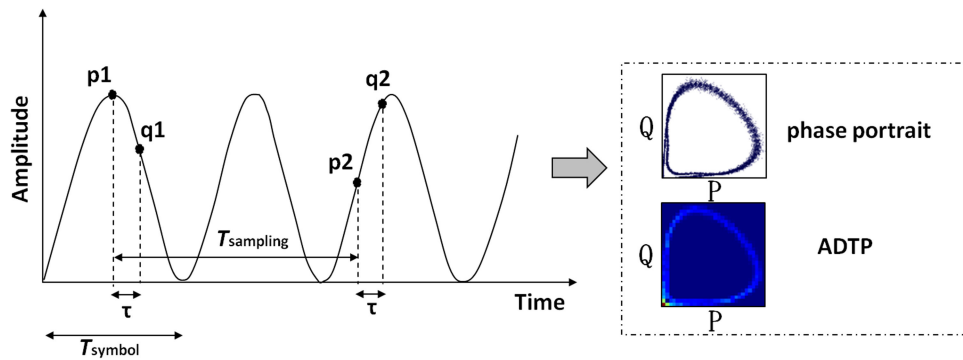


Fig. 1. The schematic diagram of the principle of ADTS.

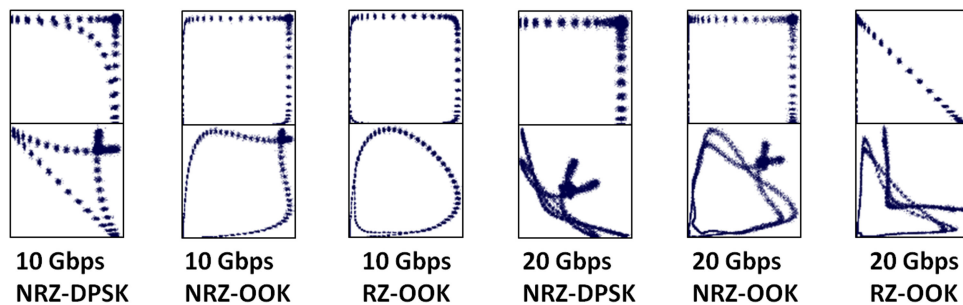


Fig. 2. Phase portraits for all six different signals for various combination of impairments. The first row corresponds to OSNR = 24 dB without CD and DGD, while the second row corresponds to OSNR = 20 dB, CD = 100 ps/nm, and D-GD = 5 ps ($\alpha = 45^\circ$). In the process of generating all images, the tap-delay time is set to 25 ps.

2. Operating Principle

2.1 Asynchronous Delay-Tap Sampling

The operating principle of ADTS is shown in Fig. 1. The amplitude information of the signal waveform after direct detection is sampled asynchronously in pairs (p_i, q_i). A fixed time delay τ interior of the sampling pair is called tap-delay. There is no relationship between the sampling period T_{sampling} and the symbol period T_{symbol} . The acquired sample pairs can be used to generate two kinds of images, one is the two-dimensional (2D) histogram, which is called ADTP, and the other is the phase portrait. Different from the conventional synchronous eye-diagrams, the generation of the ADTS-image does not require time/clock information [15], [16]. The ADTPs are often used to provide statistical features for MLAs which are limited in the ability of feature extraction. However, with the help of CNN, we can automatically discover the features from raw phase portraits. Therefore, phase portraits are used in our work. Moreover, the phase portraits of six signals (three formats with two bit rates) under various kinds of transmission impairments are shown in Fig. 2. Obviously, the patterns reflected in phase portraits are sensitive to optical impairments and signal types. Therefore, the phase portraits can provide effective features for joint OPM and BR-MFI tasks.

2.2 CNN-Based MTL Model

Convolutional network called as CNN, is used to process image data that has a known grid-like topology [17]. In CNN, two basic components are convolutional layer and pooling layer. The convolutional layer is composed of a set of kernels, which are usually a multidimensional array of parameters that are adapted by the learning algorithm. The kernel has a receptive field that embed

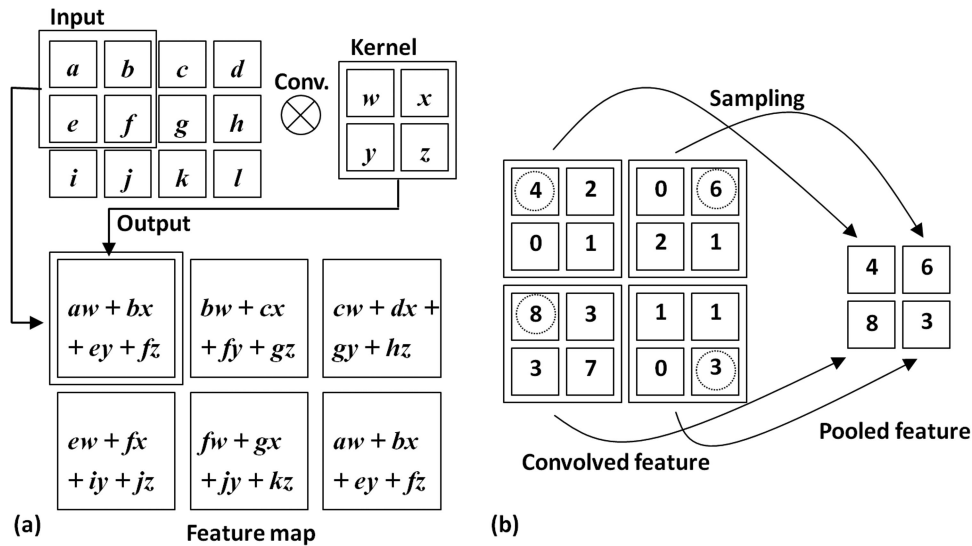


Fig. 3. Schematic diagram of (a) 2D convolution: a 3×4 input image convolution with a 2×2 kernel to generate a 2×3 feature map; (b) Max pooling: select the maximum value of each 2×2 region to generate the pooled features.

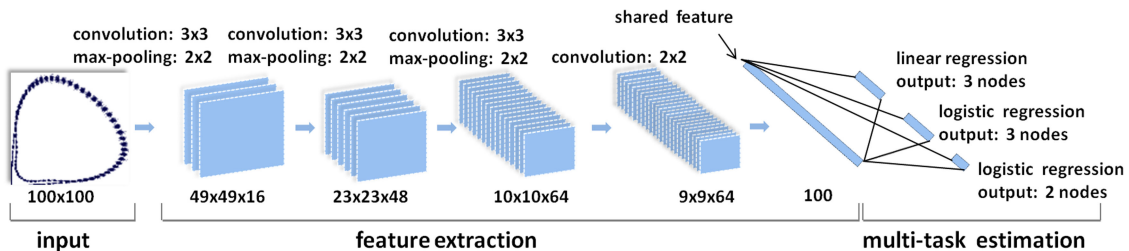


Fig. 4. Structure specification for CNN-based MTL model. Input layer: phase portraits with pixel size of 100×100 . First layer with a sixteen 3×3 convolution kernels and a 2×2 sub sampling produce sixteen 49×49 feature maps. Second layer with a forty-eight 3×3 convolution kernels and a 2×2 sub sampling produce forty-eight 23×23 feature maps. Third layer with same kernel size produce sixty-four 10×10 feature maps. Fourth layer with only a sixty-four 2×2 convolution kernels produce sixty-four 9×9 feature maps. The fifth layer with 100 nodes fully connected with fourth layer. Finally, the output layer fully connected with all 100 nodes of fifth layer and consisting of 3 nodes (linear regression task for impairments monitoring), 3 nodes (logistic regression task for modulation format identification), and 2 nodes (logistic regression task for bit-rate identification), respectively. The activation function used in the whole CNN is absolute tangent function.

to the whole depth of the input image. Fig. 3(a) shows an instance of 2D convolution. After the convolution operation of the 2×2 kernel and the 3×4 input image, a 2×3 feature map is obtained. Kernel plays a role of feature detector, detecting the same pattern everywhere of the input image then producing the feature map of 2D structure. In order to establish an effective model, multiple kernels are needed to detect multiple features in the convolution layer to generate multiple feature maps. Each feature map in the convolution layer represents the feature extracted from the input image. The pooling layer merges semantically similar features extracted from the convolution layer into one. For example, the max pooling operation reports the maximum output within a rectangular neighborhood. As shown in Fig. 3(b), the adjacent 2×2 unit region in the convolved feature map will be calculated by the pooling unit to select the maximum value as the pooled feature map. With the help of CNN, features are extracted from raw phase portraits. Moreover, a CNN-Based MTL model is used to extract and share features among different tasks. The principle of the CNN-based MTL model structure for joint OPM and BR-MFI is illustrated in Fig. 4. The input of the model is

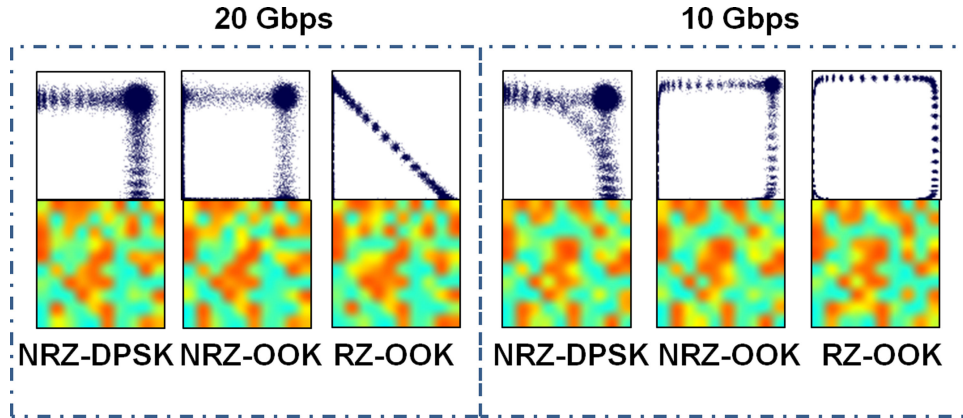


Fig. 5. The shared feature extracted from phase portraits. The first row shows the phase portraits and the second row shows the corresponding features in shared feature layer. It is obvious that phase portraits under same attribute (bit-rate or format) would produce similar feature images, which reveals that the MTL do have the ability of feature sharing.

an 100×100 gray-scale phase portrait. The feature extraction stage consists of four convolutional layers, three pooling layers and a fully connected layer. The fully connected layer called shared feature layer in the feature extraction stage produces 100 nodes which are shared by OPM and BR-MFI tasks in the estimation stage. Finally, the shared feature is fully connected with three tasks, which are comprised of one linear regression task (output 3 nodes for CD, OSNR, DGD estimation) and two logistic regression tasks (output 3 nodes for format identification and 2 nodes for bit-rate identification). Suppose we have a set of training data denoted as $\{X_i\}_{i=1}^N$ and for different tasks, their corresponding labels are $\{Y_i^{im}, Y_i^{bit}, Y_i^{mf}\}_{i=1}^N$, where Y_i^{im} is the label of impairments monitoring and the rest are the labels of BR-MFI tasks. More specifically, $Y_i^{im} \in \mathcal{R}^3$ of OSNR, CD and DGD values, $Y_i^{mf} \in \{0, 1, 2\}$ indicates three different modulation formats, and $Y_i^{bit} \in \{0, 1\}$ indicates two bit-rates. The least square and cross-entropy are used as loss functions for OPM and BR-MFI tasks respectively. Therefore, the overall loss function can be written as:

$$\text{Loss} = \text{Argmin} \left(\frac{1}{2} \sum_{i=1}^N \|Y_i^{im} - f(X_i; W^{im})\|^2 - \sum_{i=1}^N \lambda_1 Y_i^{bit} \log(p(Y_i^{bit}|X_i; W^{bit})) \right. \\ \left. - \sum_{i=1}^N \lambda_2 Y_i^{mf} \log(p(Y_i^{mf}|X_i; W^{mf})) \right)$$

$f(X_i; W^{im})$ is the OPM prediction, N is the number of training samples, and λ_1, λ_2 are the weights to balance the importance of bit-rate identification and format identification tasks. W^{im}, W^{bit}, W^{mf} represent model parameters of impairments monitoring, bit-rate identification and format identification, respectively. The overall loss function can be minimized by gradient descent method to improve the performance of the MTL model. Several phase portraits and their features (shape changed from 100×1 to 10×10) of the shared feature layer can be visualized in Fig. 5, which shows that learned features have similar patterns under same attribute.

3. System Configuration and Results

Based on VPI Transmission Maker 9.0 and Tensorflow library [18], [19], we set up a simulation system for the purpose of abstaining phase portraits as well as verifying the performance of CNN-based MTL technique. Fig. 6 shows the system configuration used in our simulation. Six different transmitters are used to generate six different signals (10/20 Gbps RZ-OOK, 10/20 Gbps NRZ-OOK

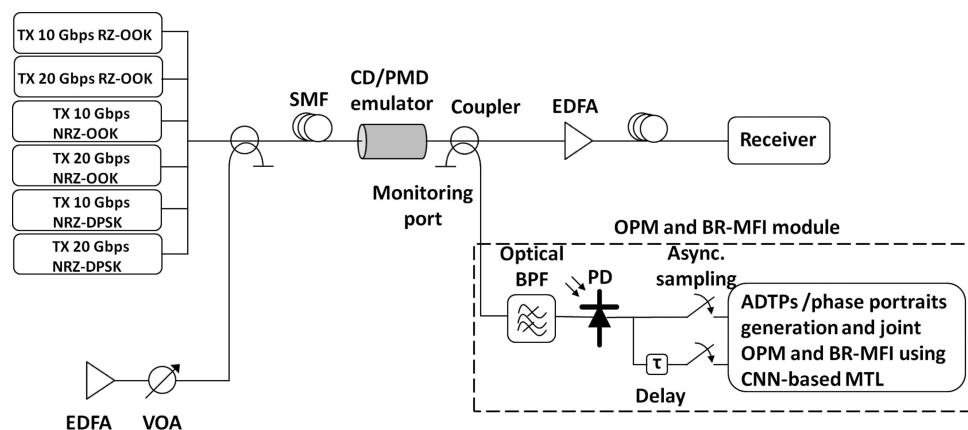


Fig. 6. Block diagram for joint OPM and BR-MFI using CNN-based MTL technique and phase portraits.

and 10/20 Gbps NRZ-DPSK). Then a single-mode fiber (SMF) is used to transmit all signals. In order to simulate real signal impairments, an Erbium-doped fiber amplifier (EDFA) is used to add amplified spontaneous emission (ASE) into the signal. Then, a variable optical attenuator (VOA) is used to adjust the values of signal's OSNR, which is varied in the range of 10–28 dB (each step is 2 dB). The CD emulator is used to adjust the accumulated CD value of the link, which is varied in the range of 0–450 ps/nm (each step is 50 ps/nm). Similarly, the PMD emulator is used to adjust DGD value to the range of 0–10 ps (each step is 1 ps). For each DGD value, the angle α between the transmitted signal's state of polarization (SOP) of the PMD emulator is altered at seven random values (0° , 15° , 30° , 45° , 60° , 75° , 90°).

The OPM and BR-MFI module is used to receive optical signals extracted from the link by an optical coupler. Inside the OPM and BR-MFI module, an optical band-pass filter (BPF) is used to filter the desired channel. Then, a receiver with optical and electrical bandwidth of 0.8 nm and 50 GHz is used to directly detect the filtered signal. After direct detection, an electrical coupler is used to divide two electrical signals from the original one. Finally, the tap-delay is added on one of the two electrical signals. The selection of the tap-delay in phase portraits has no strict constraints and often a tap-delay is a fraction of the symbol period [20], [21]. A tap-delay of 25 ps is used, since it results in phase portraits with significant differences. A phase portrait (in ".png" format) is generated by 200,00 delay-tap sample pairs. A large data set encompassing 660, 0 phase portraits corresponding to different OSNR, CD, DGD, α , bit-rates, and modulation formats is obtained. Each phase portrait has impairments, bit-rate, and format label vectors consisting of 3 values, 2 values and 3 values, respectively. The impairments label vector denote 3 values of CD, OSNR and DGD which are normalized to [0, 1]. The bit-rate label vector of 2 values denote two different bit-rates (10 Gbps: 10, 20 Gbps: 01). Similarly, 3 values of format label vector denote modulation formats (NRZ-DPSK: 100, NRZ-OOK: 010, RZ-OOK: 001). We randomly divide all the acquired phase portraits into training set and testing set according to the ratio of 90% and 10%, respectively. During training process, CNN-based MTL can extract and share features to minimize the errors between ground truth and output values through gradient descent methods. To investigate the effect of the model performance after training, the accuracies of the BR-MFI and RMS errors of impairments monitoring are measured by the testing set. The PCA-based joint OPM and BR-MFI machine learning algorithm is used to demonstrate the comparative advantage of CNN-based MTL.

3.1 Impairments Monitoring

Firstly, the influence of related tasks is examined to see whether the information of bit-rate and format is useful for impairments monitoring task. We evaluate four different models. Specifically, the first model is trained only on OPM task. Another two models are trained on OPM task along

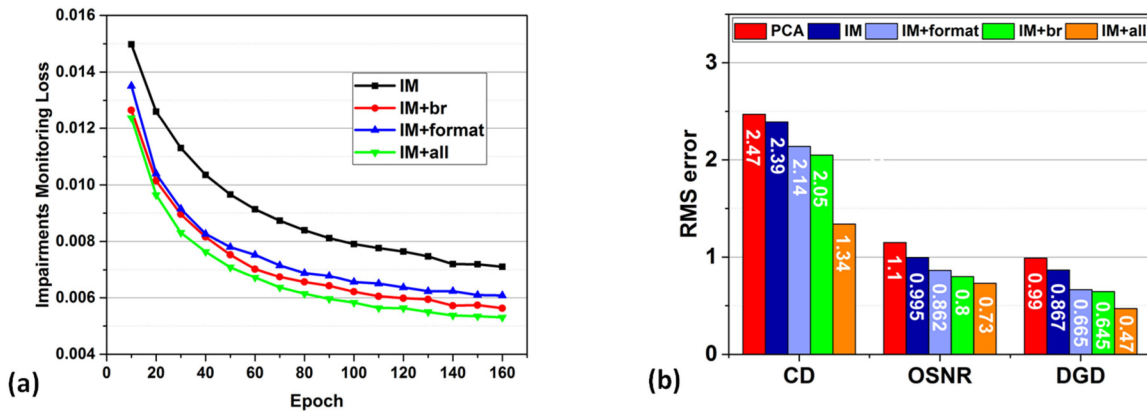


Fig. 7. (a) Impairment monitoring loss as a function of epochs for four different MTL models; (b) Performance comparison among different models and PCA algorithms. The extracted feature vector size of phase portraits by PCA is 30.

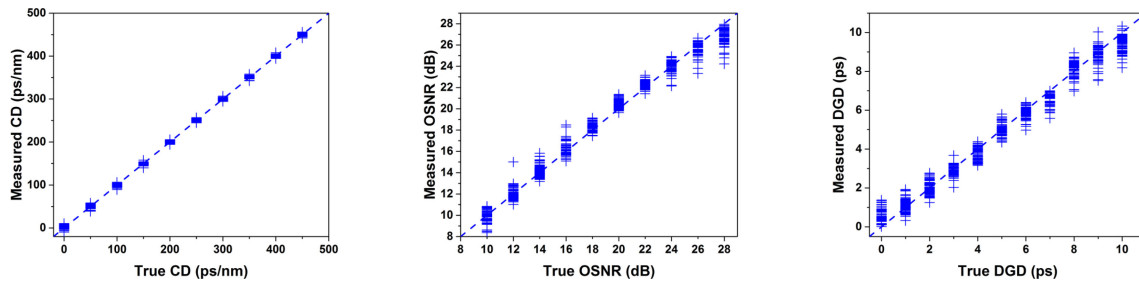


Fig. 8. Simulation results on testing set of CD (RMS error of 1.34 ps/nm), OSNR (RMS error of 0.73 dB), and DGD (RMS error of 0.47 ps) by IM+all model.

with the related tasks of bit-rate (br) identification and format identification, respectively. Finally, the whole model is trained using all the three related tasks. To do this simply, each variants are named by impairments monitoring (IM) and the related task, such as “IM” ($\lambda_1 = \lambda_2 = 0$), “IM+br” ($\lambda_1 = 1$, $\lambda_2 = 0$), “IM+format” ($\lambda_1 = 0$, $\lambda_2 = 0.4$), “IM+all” ($\lambda_1 = 0.8$, $\lambda_2 = 0.6$). Fig. 7(a) plots the impairments monitoring losses of the four MTL models, and the RMS error over different impairments is showed in Fig. 7(b). In particular, IM+all model outperforms IM model by a large margin. When single related task is present, IM+br model outperforms IM+format model in a very close way. This means under our configuration, bit-rate and format information may have similar effect on impairments monitoring task. The RMS errors of IM+all model for CD, OSNR and DGD are 1.34 ps/nm, 0.73 dB, and 0.47 ps, respectively. All four MTL models outperform the PCA algorithm significantly since PCA lacks of feature extraction and feature sharing. The results indicate that optimizing impairments monitoring with bit-rate and format identification are beneficial. Note that the accuracies of identification tasks of IM+br, IM+format, and IM+all are all 100%.

The IM+all model predicted OSNR, CD, and DGD values against the true impairment values are shown in Fig. 8. It is clearly seen that a good linear relationship is satisfied between the predict values and true values. However, when at small (CD: 0~150 ps/nm, OSNR: 10 ~ 16 dB, DGD: 0 ~ 3 ps) and large (CD: 350~450 ps/nm, OSNR: 24~28 dB, DGD: 7~10 ps) value region, the scatter points are more loose than at middle (CD: 200~300 ps/nm, OSNR: 18~22 dB, DGD: 4~6 ps) value region. The explanation of the above result is that when CD, OSNR and DGD start to increase at a small value, every step of growth will bring obvious distinguish on phase portraits, which is helpful for CNN-based MTL monitoring. However, when these values continue to increase and surpass a certain threshold, the effect on phase portrait is getting weaker and weaker. In this case, the

TABLE 1
Summary of the Monitoring Results

Signal	OSNR range (RMS error)	CD range (RMS error)	DGD range (RMS error)
10Gbps NRZ-OOK	10-28 dB (0.47 dB)	0-450 ps/nm (1.30 ps/nm)	0-10 ps (0.62 ps)
20Gbps NRZ-OOK	10-28 dB (0.60 dB)	0-450 ps/nm (1.49 ps/nm)	0-10 ps (0.44 ps)
10Gbps RZ-OOK	10-28 dB (0.53 dB)	0-450 ps/nm (1.20 ps/nm)	0-10 ps (0.53 ps)
20Gbps RZ-OOK	10-28 dB (1.17 dB)	0-450 ps/nm (1.00 ps/nm)	0-10 ps (0.30 ps)
10Gbps NRZ-DPSK	10-28 dB (0.38 dB)	0-450 ps/nm (1.53 ps/nm)	0-10 ps (0.48 ps)
20Gbps NRZ-DPSK	10-28 dB (0.98 dB)	0-450 ps/nm (1.38 ps/nm)	0-10 ps (0.37 ps)

adjacent values of small and large region present similar features, which is difficult for CNN-based MTL monitoring and increase the prediction errors. Table 1 tabulates the monitoring results of each signal. As shown in the table, the proposed technology has a wide monitoring range and acceptable monitoring precision for impairments in various combination of different bit rates and formats. In addition, the computation time for PCA-based algorithm and CNN-based MTL algorithm based on an ordinary computer (Intel Core i7 CPU) are calculated respectively. We found that as long as the feature extraction process by PCA is finished, the PCA-based pattern recognition algorithm does not need the process of training, but it takes a lot of time to process a test image (~ 1.3 s). The CNN-based MTL model spends 9 s on each epoch during training process, and the test time of each image is ~ 56 ms, which is much shorter. This means that CNN-based MTL model can be used for real-time processing as long as the training process is completed.

3.2 Bit-Rate and Modulation Format Identification

Next, we investigate model performance for BR-MFI in detail. Fig. 9(a) shows the performance of BR-MFI. The *br_only* model is trained only on bit-rate identification task. Similarly, the *format_only* model is trained only on format identification task. The accuracies of bit-rate and format from IM+all model are named as *br_IM+all* and *format_IM+all*, respectively. When the accuracies for four models of *br-only*, *br_IM+all*, *format_only*, *format_IM+all* attain 100%, the minimum epochs are 27, 34, 12, 17, respectively. It is found that when identification task and monitoring task are trained together (IM+all model), the epochs are increased compared with individual training results when accuracy equals 100%. The reason is that joint optimization of different tasks make the model extract and share useful features for all tasks simultaneously, which is more difficult and time-consuming than single identification task. Moreover, whether the identification task is trained individually or together with monitoring task, the bit-rate identification task requires more training epochs than the format identification task when the accuracy equals 100%, which illustrate formats are much easier to be recognized than bit rate. It is because different format present notable different patterns in phase portraits than bit-rate. Fig. 9(b) shows the identification performance of the PCA-based algorithm and IM+all model on all the six signals. In particular, the CNN-based MTL achieves 100%

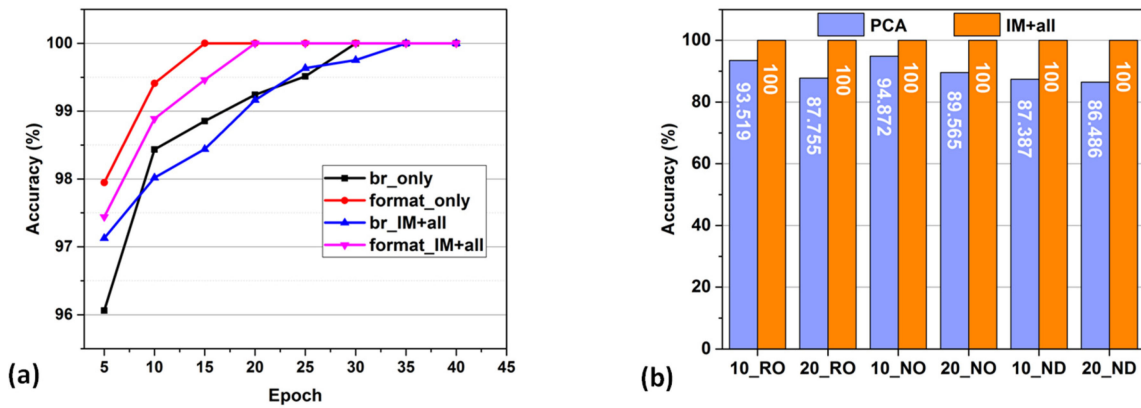


Fig. 9. (a) Accuracy of BR-MFI for three different models. The br_only curve and the format_only curve represent two models that are only trained with bit-rate identification task and format identification task, respectively. The br_IM+all curve and the format_IM+all curve represent the identification performance of IM+all model on bit-rate identification and format identification, respectively; (b) BR-MFI accuracies compared between the PCA-based algorithm and IM+all model. For the sake of simplicity, we name signal type NRZ-OOK, RZ-OOK, NRZ-DPSK as NO, RO and ND, respectively. And then put the bit-rate in front of the abbreviation such as 10_NO. The extracted feature vector size of phase portraits by PCA is 30.

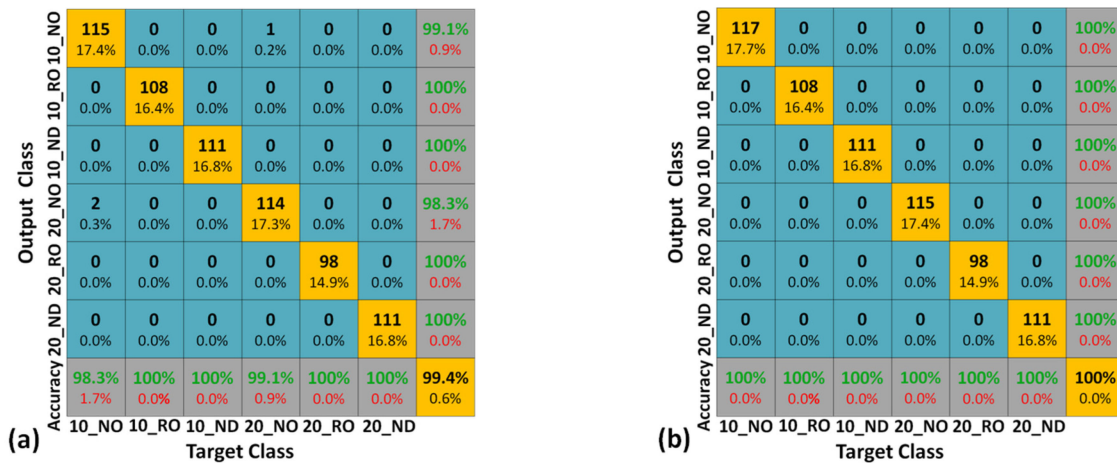


Fig. 10. Confusion matrix for BR-MFI under different size of training set, the epoch is set to 35 and the original testing set is used: (a) the number of training set is 77.8% (4620) of original training set, One 20 Gbps NRZ-OOK phase portrait is misclassified as 10 Gbps NRZ-OOK and Two 10 Gbps NRZ-OOK are misclassified as 20 Gbps NRZ-OOK; (b) original training and testing set are used and the total accuracy is 100%.

accuracy outperform the PCA-based algorithm performance, for all six signal types, as expected, demonstrating the excellent performance of CNN-based MTL on BR-MFI.

As shown in Fig. 10, the accuracies of BR-MFI in the thirty-fifth epoch are measured when the number of training data is 77.8% (4620 images) and 100% (5940 images) of the original training set respectively. Fig. 10(a) shows the confusion matrix of the reduced training data result, which has lower accuracy since the number of training samples is limited. One 20 Gbps NRZ-OOK phase portrait is misclassified as 10 Gbps NRZ-OOK and Two 10 Gbps NRZ-OOK are misclassified as 20 Gbps NRZ-OOK. As a result, the overall accuracy drop to 99.4%. Fig. 10(b) shows that when the number of training data is restored to the original size, the accuracy is restored to 100%. Further more, we continue to reduce the size of the training set and measure the accuracies of each

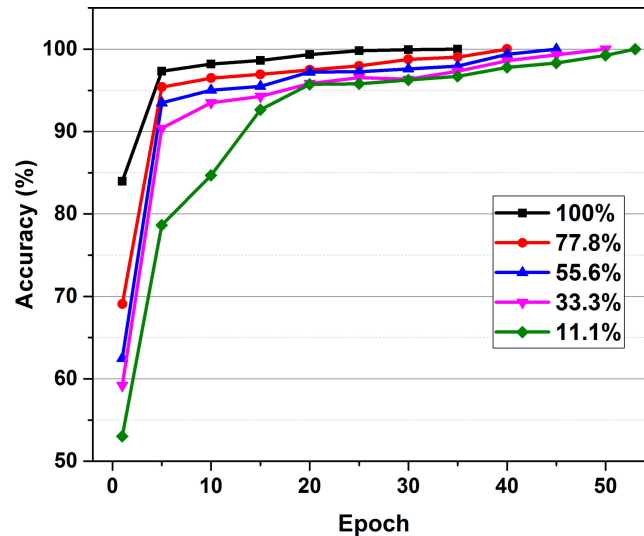


Fig. 11. The BR-MFI accuracies at different epochs for different sizes of training set. The proportion of the training set to the original one is 100%, 77.8%, 55.6%, 33.3%, and 11.1%.

epoch for the purpose of exploring the limit of CNN-based MTL identification for BR-MFI, as shown in Fig. 11.

The proportion of the training set to the original one is 100% (5940 images), 77.8% (4620 images), 55.6% (3300 images), 33.3% (1980 images), and 11.1% (660 images). Obviously, at the same epoch, the smaller the training set, the worse the performance. As the epoch increased, the accuracies eventually increased to 100% for all sizes of the data. Even the smallest size of the training set can achieve the error-free results by increasing several epochs, which shows the prominent effect of the CNN-based MTL on BR-MFI.

3.3 Task Weights and Model Structure

Then we study the influence of task weights and model structure on the performance of CNN-based MTL. Here, we focus mainly on the OPM task rather than BR-MFI task. At the very beginning, the task weights λ_1 and λ_2 are adjusted when the model structure and other parameters remain unchanged. λ_1 , λ_2 are changed from 0 to 1.2 with step of 0.2. The impairments monitoring losses under the combination of λ_1 and λ_2 are recorded in Fig. 12. It is obvious that at the middle and low range (0.0~0.7) of λ_1 , the impairments monitoring losses are pretty high for almost full range (0.0~1.2) of λ_2 . Similarly, at the top range (0.9~1.2) of λ_2 , the impairments monitoring losses are pretty high for almost full range (0.0~1.2) of λ_1 . The lower loss value can be achieved when λ_1 at the range of 0.8~1.0 and λ_2 at the range of 0.6~0.8. This demonstrates that task weights do have a significant impact on model performance. When λ_1 at top range and λ_2 at medium range, the information provided by BR-MFI tasks can help OPM task to get optimal results.

Next, the influence of the model structure on the performance of the CNN-based MTL is further studied. The structure of the model is mainly composed of the number of feature map in each convolution layer (Conv1, Conv2, Conv3, Conv4). The original model structure are (16, 48, 64, 64), and now we change the number of feature maps to construct different structures: (4, 12, 16, 16), (8, 24, 32, 32), (32, 96, 128, 128), (64, 192, 256, 256). The impairments monitoring losses at different epochs are measured based on different model structures, as shown in Fig. 13. It is seen that when the scale of the model increased gradually from (4, 12, 16, 16) to (32, 96, 128, 128), the loss is getting smaller and smaller. However, when the network structure continues to grow to (64, 192, 256, 256), the loss increased suddenly. The large-scale structures like (64, 192, 256, 256) makes

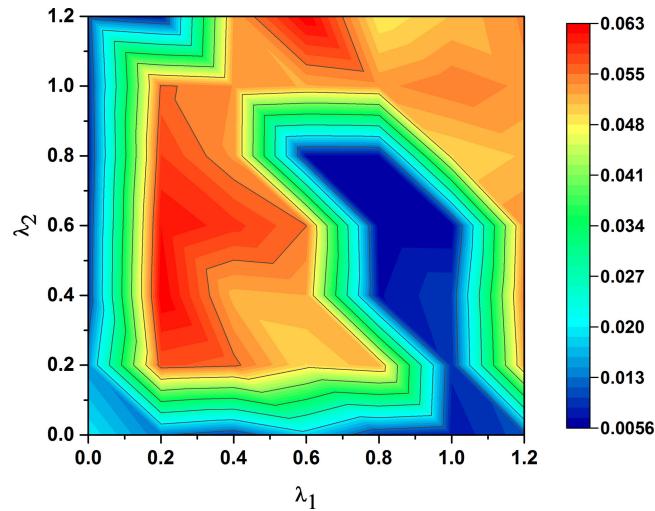


Fig. 12. The heatmap of impairments monitoring losses under different task weights values of λ_1 and λ_2 . The warmer the color, the greater the monitoring loss.

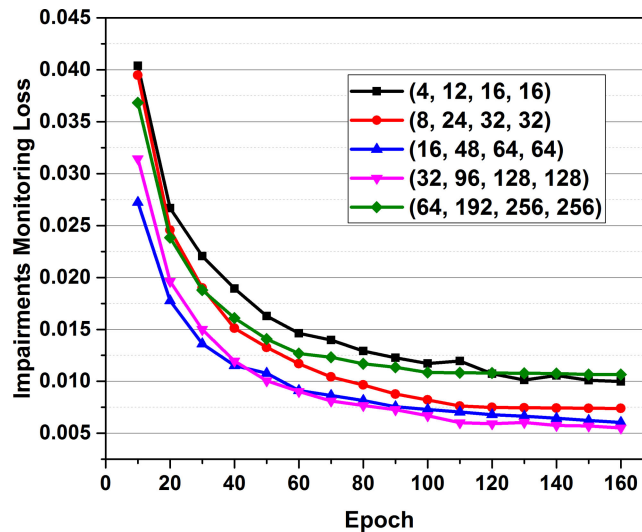


Fig. 13. The impairments monitoring losses at different epochs with different model structures of (4, 12, 16, 16), (8, 24, 32, 32), (16, 48, 64, 64), (32, 96, 128, 128), (64, 192, 256, 256).

the number of model parameters increase sharply, resulting in deterioration of model performance and cannot detect more useful features.

As analyzed above, we can conclude that it is crucial to select suitable task weights and design reasonable model structure for specific situation. For phase portrait analysis, the task weights of λ_1 , λ_2 set to 0.8, 0.6 respectively and structure of (16, 48, 64, 64) is a reasonable and practical option.

4. Conclusion

In summary, a CNN-based MTL algorithm was proposed to joint optimization of OPM and BR-MFI. With the ability of feature extraction and feature sharing, the performance of OPM has been improved without the reduction of BR-MFI performance. Without any prior information, the proposed CNN-based MTL technique can joint detect channel impairments, bit rates and modulation formats.

Compared with PCA-based pattern recognition algorithm, CNN-based MTL performed the better results. Numerical simulation results demonstrate that the monitoring of OSNR, CD, and DGD with RMS error of 0.73 dB, 1.34 ps/nm, and 0.47 ps, respectively. Similarly, the accuracy of BR-MFI achieve 100% even with limited size of training data. Moreover, the effects of training data size, task weights and model structure on CNN-based MTL performance were separately studied. Our algorithm is efficient because it allows real-time performance (~ 56 ms) on Intel Core i7 CPU without the usage of Graphics Processing Unit (GPU) and other acceleration devices. This technique has the potential to perform intelligent signal analysis for future heterogeneous optic communication networks.

References

- [1] Z. Pan, C. Yu, and A. E. Willner, "Optical performance monitoring for the next generation optical communication networks," *Opt. Fiber Technol.*, vol. 16, no. 1, pp. 20–45, 2010.
- [2] 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," *Photon. Technol. Lett.*, vol. 24, no. 12, pp. 982–984, 2012.
- [3] F. N. Khan *et al.*, "Non-data-aided joint bit-rate and modulation format identification for next-generation heterogeneous optical networks," *Opt. Fiber Technol.*, vol. 20, no. 2, pp. 68–74, 2014.
- [4] D. Wang *et al.*, "Nonlinearity mitigation using a machine learning detector based on k-Nearest neighbors," *Photon. Technol. Lett.*, vol. 28, no. 19, pp. 2102–2105, 2016.
- [5] D. Wang *et al.*, "Combatting nonlinear phase noise in coherent optical systems with an optimized decision processor based on machine learning," *Opt. Commun.*, vol. 369, pp. 199–208, 2016.
- [6] M. C. Tan, F. N. Khan, W. H. Al-Arashi, Y. Zhou, and A. P. T. Lau, "Simultaneous optical performance monitoring and modulation format/bit-rate identification using principal component analysis," *J. Opt. Commun. Netw.*, vol. 6, no. 5, pp. 441–448, 2014.
- [7] F. N. Khan *et al.*, "Experimental demonstration of joint OSNR monitoring and modulation format identification using asynchronous single channel sampling," *Opt. Exp.*, vol. 23, no. 23, pp. 30337–30346, 2015.
- [8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [9] W. Liu, T. Mei, Y. Zhang, C. Che, and J. Luo, "Multi-task deep visual-semantic embedding for video thumbnail selection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3707–3715.
- [10] E. Parisotto, J. L. Ba, and R. Salakhutdinov, "Actor-mimic: Deep multitask and transfer reinforcement learning," *Comput. Sci.*, 2015.
- [11] L. Zhao, Q. Sun, J. Ye, F. Chen, C. Lu, and N. Ramakrishnan, "Feature constrained multi-task learning models for spatiotemporal event forecasting," *Trans. Knowl. Data Eng.*, vol. 29, no. 5, pp. 1059–1072, 2017.
- [12] Z. Wu, C. Valentini-Botinhao, O. Watts, and S. King, "Deep neural networks employing multi-task learning and stacked bottleneck features for speech synthesis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2015, pp. 4460–4464.
- [13] Y. Zhang and Q. Yang, "An overview of multi-task learning," *Nat. Sci. Rev.*, vol. 5, no. 1, pp. 30–43, 2017.
- [14] D. Wang *et al.*, "Modulation format recognition and OSNR estimation using CNN-Based deep learning," *Photon. Technol. Lett.*, vol. 29, no. 19, pp. 1667–1670, 2017.
- [15] S. D. Dods and T. B. Anderson, "Optical performance monitoring technique using delay tap asynchronous waveform sampling," in *Proc. Opt. Fiber Commun. Conf. Nat. Fiber Optic Eng. Conf.*, 2006, pp. 3pp.
- [16] F. N. Khan, A. P. T. Lau, Z. Li, C. Lu, and P. K. A. Wai, "Statistical analysis of optical signal-to-noise ratio monitoring using delay-tap sampling," *Photon. Technol. Lett.*, vol. 22, no. 3, pp. 149–151, 2010.
- [17] I. Goodfellow *et al.*, *Deep Learning*. Cambridge: MIT Press, 2016.
- [18] VPIsystems, "VPItransmissionMaker," Somerset, 2009.
- [19] M. Abadi, "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," in *Proc. Conf. Lang. Resources. Eval.*, 2016, pp. 3243–3249.
- [20] B. Kozicki, A. Maruta, and K. I. Kitayama, "Experimental investigation of delay-tap sampling technique for online monitoring of RZ-DQPSK signals," *Photon. Technol. Lett.*, vol. 21, no. 3, pp. 179–181, 2009.
- [21] B. Kozicki, A. Maruta, and K. Kitayama, "Transparent performance monitoring of RZ-DQPSK systems employing delay-tap sampling," *J. Opt. Netw.*, vol. 6, no. 11, pp. 1257–1269, 2007.