PAN

Opto-Electronics Review

journal homepage: https://journals.pan.pl/opelre

Using ADTS and CNN methods to effectively monitor CD, crosstalk, and OSNR in an optical network

Tomasz Mrozek*, Krzysztof Perlicki

Institute of Telecommunications, Warsaw University of Technology, ul. Nowowiejska 15/19, 00-665 Warsaw, Poland

Article info

Article history:
Received 08 Feb. 2024
Accepted 25 Jul. 2024
Available on-line 22 Sep. 2024

Keywords:
optical performance monitoring;
asynchronous delay-tap sampling;
convolutional neural network;
multi impairments monitoring.

Abstract

The article presents the results of a method based on asynchronous delay-tap sampling (ADTS) and convolutional neural network (CNN) for determining simultaneously occurring disturbances described using the chromatic dispersion (CD), crosstalk and optical signal-to-noise ratio (OSNR) parameters. The ADTS method was used to generate training and test data for the convolutional network, which in turn was used to learn to recognize interference from said data. The tests were carried out for a transmission speed of 10 Gbit/s and for on-off keying (OOK) and differential phase shift keying (DPSK) modulation. Very good results were obtained in recognizing simultaneously occurring phenomena. Accuracy of over 99% was achieved for CD and crosstalk for DPSK modulation and over 98% for OOK modulation. In the case of amplified spontaneous emission (ASE) noise, slightly weaker results were obtained, above 95–96% for both modulations. Based on the conducted research, it was determined that the use of ADTS and CNN methods enables monitoring of simultaneously occurring CD, crosstalk, and ASE noise in the physical layer of the optical network, while maintaining the requirements for modern monitoring systems.

1. Introduction

To meet contemporary requirements for the implementation of new generation multimedia services requiring high-data transmission speeds, systems are used that include network elements such as reconfigurable optical add-drop multiplexers (ROADM) and optical cross-over frames – optical cross connect (OXC), as well as tunable lasers and filters. These systems are based on increasingly complex mesh topology or mixed ring-mesh topology. To ensure the greatest possible efficiency of optical links, additional technical solutions are used to increase the information capacity (its throughput). One of the most important is the multiplexing technique. The most frequently techniques used are wavelength division multiplexing (WDM) and time division multiplexing (TDM), which are implemented in long-range networks, metropolitan and access networks. The implementation of such advanced transmission systems introduces an additional level of complexity in the physical layer of the optical network and reduces the transparency in its management process, and increases the susceptibility to interference, limiting the system ability to transmit information. Various phenomena may affect the functioning of the network and errors during transmission (e.g., attenuation, chromatic dispersion (CD), amplified spontaneous emission noise, crosstalk, etc.). By monitoring the appropriate parameters of the transmitted signal, it is possible to assess the quality of transmission and the intensity of the impact of these phenomena.

With a development of high-speed optical networks, multi-channel transmissions and new modulation formats were introduced and the requirements for rapid network reconfiguration also increased. Additionally, there has been an increase in the demand for universal and fast methods for assessing the quality of optical transmission which would allow obtaining (using one method) as much information as possible about the many possible causes of deterioration of transmission quality. For this reason, techniques enabling simultaneous measurement of several signal parameters using one method have become a very important direction in the development of monitoring the quality of optical networks. Modern methods of optical performance monitoring (OPM) require high accuracy and

*Corresponding author at: tomasz.mrozek@pw.edu.pl

https://doi.org/10.24425/opelre.2024.151963
1896-3757/ Association of Polish Electrical Engineers (SEP) and Polish Academic of Sciences (PAS). Published by PAS

a large measurement range, quick measurement time, and determination of the parameters of the measured quantities, regardless of the occurring interferences or other unfavourable measurement conditions. Another very important requirement is the functionality that allows simultaneous determination of the values of many parameters, regardless of the transmission speed and modulation format. Moreover, such techniques should be characterised by low system complexity, low energy consumption and the possibility of being used for real-time measurements.

To meet the above-mentioned requirements, it was decided to use the asynchronous delay-tap sampling (ADTS) method which allows for quick measurement and graphical presentation of the signal waveform considering simultaneously occurring interferences. The graphical form with disturbances must then be properly interpreted using an appropriate technique to determine the numerical values of these disturbances. For this reason, it was decided to use convolutional neural networks (CNN), whose main field of application are issues related to image analysis and processing.

2. Asynchronous delay-tap sampling

The ADTS technique allows direct measurement of signal noise without the need to recover clock synchronization [1–3]. The method can be used to monitor the phenomena of chromatic and polarization dispersion, amplified spontaneous emission (ASE) noise, and optical crosstalk. Additionally, the method can be used for various modulation formats and transmission rates. A sample demonstration of the ADTS method can be found in Refs. 1–3. For 10 Gbit/s return-to-zero differential phase shift keying (RZ-DPSK) and non-return-to-zero differential phase shift keying (NRZ-DPSK) transmission, CD in the range of −720 to 800 ps/nm, differential group delay (DGD) parameter from 0 to 40 ps, optical signal-to-noise ratio (OSNR) from 15 to 35 dB, and optical crosstalk at −25 dB were measured. A schematic diagram of the ADTS measurement method along with a time course illustrating the method of signal sampling is presented in Fig. 1. After detection, the signal is divided into two transmission lines in the splitter. One of the lines introduces a set delay Δt in a controlled manner. In the last stage, the signal is sampled. A two-dimensional scatter plot is created from the collected pairs of samples (x, y) as phase portrait. The shape of the phase portrait depends on many signal parameters. If the signal waveform is undisturbed during transmission, the portrait shape takes on a simple and legible form.

Figure 2 shows example phase portraits for interference-free signals for DPSK and on-off keying (OOK) modulation and bit delays: 0.25, 0.5, and 1.

If disturbances occur, the shape of the phase portrait undergoes various distortions depending on the type and intensity of the phenomenon. The high sensitivity of the shape to disturbances means that the ADTS method can be used to monitor many phenomena simultaneously. As already mentioned at the beginning of the article, the phenomena of CD, optical crosstalk, and ASE noise were selected for the monitoring process. The ADTS method enables their monitoring in a wide range of values regardless of the transmission speed and modulation format. Each

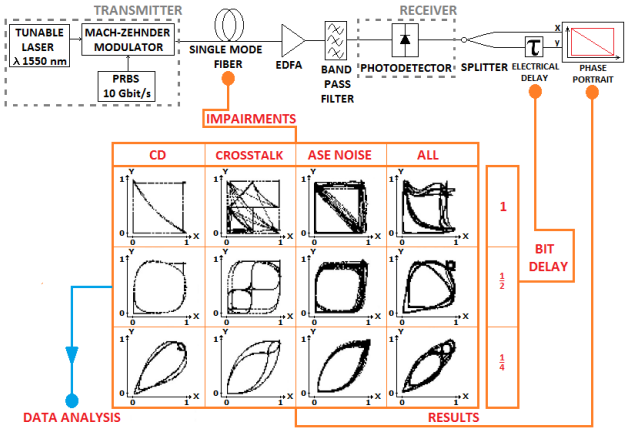

Fig. 1. ADTS – measurement model used to create phase portraits.

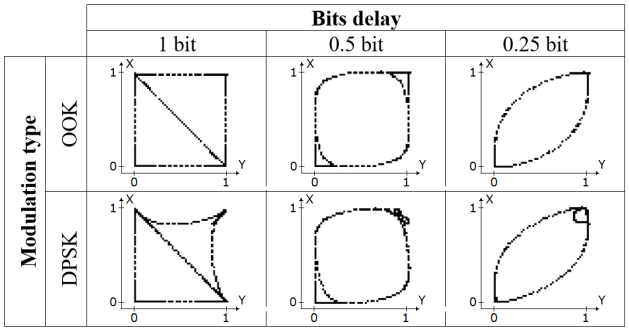

Fig. 2. ADTS – simulation model and impairments in physical layer.

phenomenon has a different impact on the graphical representation of the phase portrait of the ADTS method, and this may potentially translate into better efficiency, e.g., in the use of CNN to determine the numerical values of these phenomena. The above-mentioned phenomena can be easily modelled in the VPIphotonics simulation environment [4], thanks to which it is possible to generate a very large amount of data, which in turn is necessary for convolutional networks in the training process. Figures 3–5 show an example of the phase portrait characteristics in response to the tested disturbances for bit delay 1 and OOK modulation. When the intensity increases, each phenomenon affects a different part of the phase portrait, causing greater distortion only in this specific area.

The occurrence of CD in the transmitted signal is reflected in the phase portrait by waving the diagonal and the upper and right sides (red). At low power of the phenomenon, a clearly bent diagonal is observed. As the intensity of the phenomenon increases, the upper and right sides of the phase portrait become wavier. When the phenomenon is very intense, the shape becomes clearly stratified and its readability becomes limited.

The occurrence of low-power crosstalk is manifested by the scattering of the signal on the diagonal and slight waving on the upper and right sides of the phase portrait. The increase in the power of the crosstalk phenomenon causes the creation of additional diagonals and smaller windows, which are marked with red sections in the phase portrait.

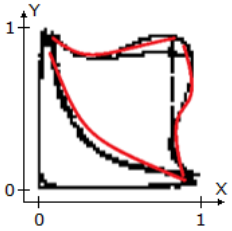

Fig. 3. The influence of CD on the phase portrait for 1 bit delay and OOK modulation.

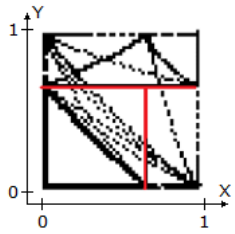

Fig. 4. The influence of intra-channel optical crosstalk on the phase portrait for 1 bit delay and OOK modulation.

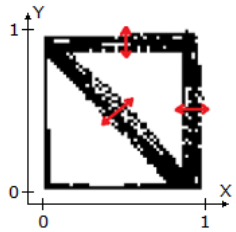

Fig. 5. The influence of ASE noise on the phase portrait for 1 bit delay and OOK modulation.

The occurrence of interference caused by the ASE noise phenomenon is manifested by signal broadening (blurring of samples) on the diagonal and on the upper and right sides of the phase portrait (red arrows). The increase in noise causes the OSNR value to decrease and the signal to become more and more scattered, until an unreadable blurred point cloud is obtained.

The ADTS technique is very good at distinguishing phenomena through graphical representation, but to determine the numerical values of individual disturbances, it is necessary to use an additional technique. To determine the values of parameters describing phenomena degrading data transmission, the following are used [5]: support vector machine (SVM), homodyne detection, Hausdorff metric, artificial neural networks, kernel function, Hough transform. Table 1 presents a summary of the mentioned methods for determining numerical values from data obtained by the ADTS method. In all works, the authors presented information on the scope of the monitored phenomena. However, some of the works omitted information about errors in determining the values of parameters describing disturbances. For these items, the "Error" field has been left blank.

The techniques used to determine numerical values so far do not meet the criteria required for modern monitoring systems. Each of them is characterised by certain limitations in the accuracy of value determination, detection range, and modulation formats (there is no information about limitations in the selection of modulation formats). The authors

Table 1.
Comparison of techniques for determining numerical values from the ADTS method.

Modulation	Impairments	Error
Support vector machine [2, 3, 6]		
10 Gbit/s NRZ-OOK, 40 Gbit/s NRZ-DPSK, 40 Gbit/s RZ-DQPSK, 80 Gbit/s PM-DQPSK	OSNR 10 – 30 dB CD −1400 – 1400 ps/nm PMD* 0 – 60 ps Crosstalk 10 – 25 dB	–
Homodyne detection [7]		
10 Gbit/s NRZ-DPSK	CD 0 – 750 ps/nm OSNR 10 – 30 dB	–
Hausdorff interval [8]		
10 Gbit/s NRZ-OOK	CD 0 – 160 ps/nm OSNR > 30 dB PMD < 15 ps	CD < 15 ps/nm, where OSNR > 30 dB and DGD < 15 ps, OSNR < 0.5 dB, where CD < 160 ps/nm and DGD < 15 ps
40 Gbit/s NRZ-DQPSK	CD 0 – 400 ps/nm	–
Artificial neural networks [9]		
10 Gbit/s NRZ-OOK	OSNR 15 – 30 dB CD 0 – 55 ps/nm PMD 0 – 30 ps	CD ± 30 ps/nm, where CD 100 – 500 ps/nm OSNR ± 2 dB, where OSNR 18 – 30 dB; DGD ± 3 ps, where DGD 5 – 25 ps
Kernel method [3]		
40 Gbit/s NRZ-DPSK	CD 0 – 400 ps/nm OSRN 15 – 25 dB PMD 0 – 22.5 ps	CD ± 11 ps/nm, where CD 0 – 400 ps/nm and OSNR 15 – 25 dB, DGD ± −1.9 ps, where DGD 0 – 22.5 ps
Hough transform [10–12]		
10 Gbit/s NRZ-DQPSK	CD 0 – 600 ps/nm OSNR 15 – 30 dB	CD ± 25 ps/nm OSNR ± 3 dB
10 Gbit/s RZ-DPSK, 20 Gbit/s RZ-DQPSK 40 Gbit/s NRZ-DQPSK	OSNR 8.7 – 35 dB CD −600 – 600 ps/nm	–

*PMD - polarization mode dispersion

of some works do not additionally provide the error of the determined values. In the case of some methods, the accuracy of determining the values of individual phenomena depends on the intensity level of co-occurring phenomena. This means that the phenomena may overlap, which can make it impossible to determine their values with appropriate accuracy. It is therefore justified to look for a new tool that will unambiguously automatically determine the parameter values of several simultaneously occurring phenomena, regardless of the bit rate, type of

modulation, and intensity of these phenomena and their interaction with each other. For this reason, it was decided to use CNN, which are presented in the next chapter.

3. Convolutional neural network

Convolutional neural networks (CNNs) or convolutional networks (ConvNet) are multi-layer neural networks used in the deep learning method, which is one of the basic machine learning techniques. The main distinguishing element of this type of networks are convolution layers which enable hierarchical feature extraction. This means that in the initial layers, the network learns to recognize basic shapes, e.g., edges, lines, dots, colours, gradients, and in the next more complex shapes, e.g., nose, eye, mouth, and in the last layers whole objects, e.g., face, car. The more hidden layers are used, the more complex features will be extracted. Convolutional layers are additionally characterised by high resistance to overfitting, i.e., a condition in which, while training the network, it is overfitted to secondary and irrelevant details that are not important from the point of view of the important features of the problem being solved. ConvNet are an example of deep neural networks. The most important scientific work that defined the basic architecture and the CNN learning algorithm is the publication by LeCun from 1998 [13]. Due to the topics discussed in this article, an example of a convolution operation for linear filtering is presented below. This is one of the most powerful image quality improvement methods in which part of the signal frequency spectrum is modified by the filter transfer function. The sum of convolutions between the input image and the filter impulse response characterises the output images. Linear image filtering is expressed in (1) [14]:

$$h(m,n) = \sum_{s=0}^{M}\sum_{t=0}^{N} g(m-s,n-t)f(s,t) = g(m,n) ** f(m,n), \tag{1}$$

where M and N – the image dimensions in pixels, $h(m,n)$ – the output image, $g(m,n)$ – the filter transfer function, $f(m,n)$ – the input image, $**$ – the two-dimensional convolution.

A typical ConvNet used in the field of image recognition is composed of an input layer, convolutional layers, reduction layers, fully connected layers, which constitute a classic neural network, and an output layer. There are three types of hidden layers in convolutional networks. Convolutional layers in which a selected activation function is performed on the convolution values. The activation function introduces a higher degree of complexity and helps the network learn complex patterns. One of the most popular activation functions is a rectifier linear unit (ReLU). Reduction layers are also known as combining or sampling layers, in which a matrix of larger pixels is converted into a matrix of smaller pixels. Fully connected layers in which a sigmoidal activation function is performed on the weighted sum of excitations, most often it is a different activation function than that used in convolutional layers, e.g., the hyperbolic tangent (tanh) function. Figure 6 shows the standard and generalized architecture of the Le-Net5 convolutional network with three hidden layers [15].

3.1. *Using CNN to monitor simultaneously occurring phenomena*

During own research, the gradual appearance of new articles dealing with similar topics (use of machine learning with ADTS in OPM) was observed [16–21]. These works focused on different modulation formats (QPSK, OOK, PAM4, PAM8, DPSK, 8QAM, 16QAM, and others) and different transmission rates (10, 20, 30, 60 Gb/s, and others). Some of the works additionally presented the use of other techniques such as: eye diagram [19], diagram analyser [18], and asynchronous amplitude histogram [17]. Only two studies tested the simultaneous monitoring of three parameters [20, 21]. In both studies, DGD was measured as one of the parameters, but this parameter had minimal impact on the shape of the phase portrait and the presented measurement range was very narrow. In the first study, the following parameters were measured: CD in the measurement range of 0–450 ps/nm (50 ps/nm jump); OSNR 10–28 dB (2 dB jump), and DGD 0–10 ps (1 ps jump). The following average accuracy was obtained: 1.52 ps/nm for CD, 0.81 dB for OSNR, and 0.32 ps for DGD. The research was performed on a data set consisting of only 1100 images for the modulations and bit rates studied. CNN models were trained over a range of 60 to 125 epochs. For the second studies found, the same three parameters were also measured. In this case, however, other modulations and transmission rates were tested. The following average accuracy was obtained: 1.34 ps/nm for CD, 0.73 dB for OSNR, and 0.47 ps for DGD. For this study, an even smaller number of images (990) included in each dataset were used. Additionally, CNNs were trained over a range of 35 to 50 epochs.

The remaining mentioned works [17, 19, 22] focused more on the simultaneous recognition of transmission rate and modulation format along with the OSNR phenomenon. Additionally, in these studies, to increase the reality of the simulation, the CD phenomenon was introduced into the optical path, but its impact on other parameters was not measured.

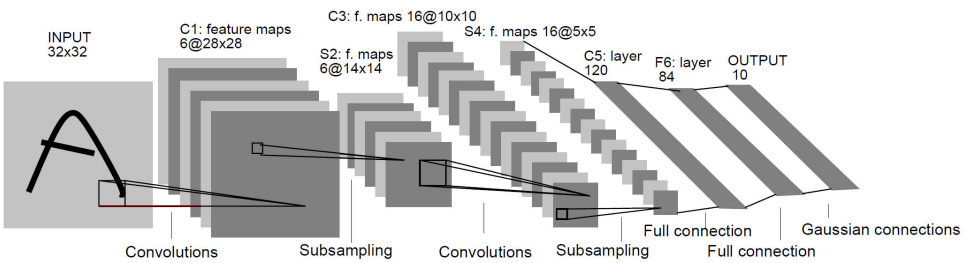

Fig. 6. Standard Le-Net5 convolutional network architecture [13] introduced by LeCun.

4. Training CNN using ADTS data

The CNN training process using ADTS data was divided into two stages. In the first part, it was necessary to prepare data for training. The data were prepared using the ADTS method modelled in the VPIphotonics software. The obtained data were prepared into special datasets. The datasets were then used as input to the built CNN using the TensorFlow library and the Python programming language. Backward linear regression was used to determine numerical values from the trained networks. Below are more detailed information on the steps listed.

4.1. Data preprocessing

To obtain data for the training process of the ConvNet, it was necessary to model a measurement station carrying out measurements in accordance with the ADTS method. To obtain good fitting results for the trained models, in addition to selecting the appropriate architecture for the network being built, it was also necessary to provide the appropriate amount of data in the training sets. Depending on the problem being solved, this number should be at least several tens of thousands. Additionally, the data should have a relatively even distribution of the values of the analysed phenomenon in relation to the investigated disturbance range. Due to the need to meet the quantitative criterion of the number of data and due to the high flexibility in generating a combination of CD, crosstalk, and ASE noise phenomena, it was decided to use a dedicated simulation environment in the form of the VPIphotonics software [4]. As a result of the simulation, data sets containing from 40 000 to 102 500 phase portraits were obtained, with a resolution of 50×50 pixels. The research showed that above the number 102 500, no better results are obtained in the process of training the CNN. Ultimately, it was decided to use data sets of 62 000 images. Data were generated in the ranges of 0–2000 ps/nm for CD and 5–40 dB for crosstalk and ASE noise. Additional parameters adopted during the simulation are bit delays of 0.25, 0.5, and 1 bit, NRZ-OOK and NRZ-DPSK modulation, bit rate of 10 Gbit/s. The selected modulations with a bit rate of 10 Gbit/s are among the basic ones used in optical transmission systems. There are many results available in the literature for these modulations and bit rates, so they constitute good comparative material for the results obtained from independent studies. The data generated in this way were further used in the process of training a CNN. Figure 7 shows example phase portraits obtained in the VPIphotonics environment.

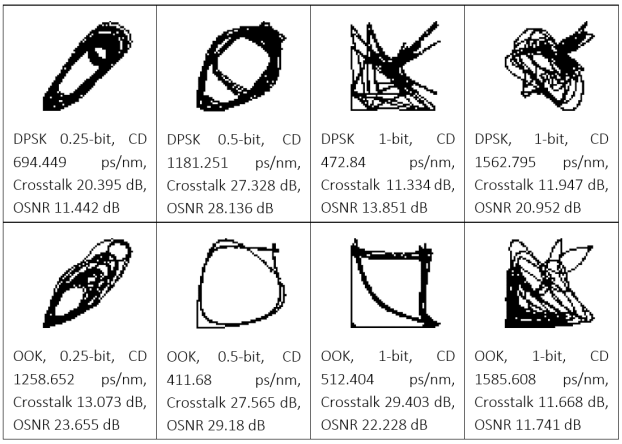

Fig. 7. Example phase portraits obtained in the VPIphotonics environment.

4.2. Building a CNN and learning process

The CNN was implemented based on the TensorFlow [23] and TFLearn [24] libraries. TensorFlow is an open-source programming library used in machine learning, including deep neural networks. The main advantage of the library is the ability to work with graphic cards. By using the computing power of the cards, it is possible to significantly accelerate calculations in the network training process. The TFLearn library is built on top of TensorFlow. It allows to quickly build convolutional networks, train them and test them. Before starting the actual training process, several network prototypes were built to assess the impact of their architecture on the learning results. Ultimately, after conducting several dozen series of tests of the training process for networks with two, three, and four convolutional layers, the one with two convolutional layers was the optimal network architecture for the considered problem. A larger number of convolutional layers significantly extended the time needed to train the network and at the same time lowered the fitness rate of the trained model, without increasing the accuracy of their operation. Figure 8 shows the architecture of the constructed convolutional network.

The first element of the network are input data sets consisting of 62 000 binary images with a resolution of 50×50 pixels. Before starting the network training process, the data are divided by the algorithm into a training set and a test set (in a ratio of 85% to 15%). The next elements of the network are two convolutional layers, which in the nomenclature of convolutional networks are referred to as hidden (deep) layers. The first convolutional

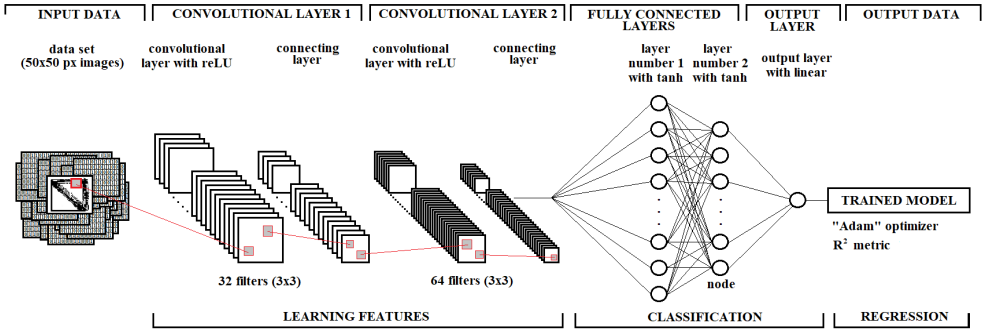

Fig. 8. Architecture of the constructed CNN.

layer is composed of 32 3 × 3 convolutional filters, while the second one is composed of 64 filters with the same dimensions as the first one. The ReLU activation function was used in each convolutional layer. The final element of each convolutional layer is the connecting layer. In the first combining layer, the dimensions of the image matrix are reduced from 50 × 50 pixels to 25 × 25 pixels and passed as input to the second convolutional layer. In the second pooling layer, the 25 × 25 pixels matrix is reduced to 13 × 13 pixels. The matrix is then flattened to one dimension and passed as input to the fully connected layers. The network architecture uses two fully connected layers that use tanh as the activation function with seven different node combinations: 1) 384 and 128; 2) 512 and 256; 3) 512 and 128; 4) 384 and 256; 5) 384 and 64; 6) 384 and 384; 7) 512 and 512 (check "Set number of fully connected layer" – Figs. 5–7). These two layers are also considered hidden layers. The last node, together with the linear activation function, constitutes the output layer of the network and determines the result of image classification into a particular interference class. However, the research conducted did not focus on classification problems, but on determining the parameter values of the occurring phenomena. For this reason, one more regression layer was used, which uses the adaptive moment estimation (ADAM) optimizer and the R^2 metric, based on which the goodness of fit of the trained model is assessed. During the empirical research, the optimal values for the parameters were also determined: snapshot step value = 1000, batch size = 100, epochs from 50 to 1000, and learning rate = 0.0001. Equation (2) presents a description of the constructed convolutional network according to the methodology available in Ref. 25.

$$Input_{image\ 2d\ (50x50)} \rightarrow \underbrace{\mathbb{C}_{3x3}^{32}\ \mathbb{P}\ \mathbb{C}_{3x3}^{64}\ \mathbb{P}}_{relu\quad relu}\ \underbrace{\mathbb{FC}^{n_1}\ \mathbb{FC}^{n_2}\ \mathbb{FC}^{out}}_{tanh\ \ tanh\ \ linear}\ \underbrace{\mathbb{R}}_{\text{"Adam", } R^2}, \tag{2}$$

where $\mathbb{C}$ is the convolution, $\mathbb{P}$ is the pooling, $\mathbb{FC}$ is the fully connected (n_1 and n_2 sequentially number of nodes in the first and second layer), and $\mathbb{R}$ is the regression. The first part of the equation (first buckle) is responsible for feature extraction, the second for classification, and the last one for regression.

The process of training the convolutional network was carried out according to the supervised learning methodology. For each phase portrait from the training data set, the given (expected) output value was known. The role of the learning algorithm was to reproduce the model fitting function as accurately as possible (with the smallest possible error) to the expected output data. Many attempts have been made to train the network for various architectures and settings to ultimately achieve network models with the best fit.

4.3. Learning results

As a result of the training process, trained CNN models were obtained. The R^2 score (coefficient of determination) was used to assess the quality of the trained models. This indicator is one of the basic ones for assessing the quality of fit of the learned model. The indicator ranges from 0 to 1. An indicator ranging from 0.9 to 1 means a very good fit. The closer to 1.0, the better the fit of the trained model. The R^2 score is described by the following (3):

$$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{3}$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} y_i \epsilon_i^2$ (source: R^2 score, the coefficient of determination 2021).

Tables 2–4 include the average error values of the determined parameters (relative to the reference values) for the highest R^2 coefficients and R^2 coefficients for the lowest average error values obtained for individual phenomena in the process of training the convolutional network.

Table 2.
Summary of the lowest average values of CD errors and R^2 fit coefficients obtained in the network training process.

Bit delay	Modulation	Average CD error / R^2	Number of epochs	Serial number of fully connected layer
average values of chromatic dispersion errors for the highest R^2 coefficients				
0.25	DPSK	1.3155 / 0.999988	650	1
	OOK	1.6099 / 0.999976		4
0.5	DPSK	1.2408 / 0.999989		1
	OOK	1.4611 / 0.999980		2
1	DPSK	1.4100 / 0.999987	250	6
	OOK	1.6441 / 0.999977	650	2
R^2 coefficients for the smallest average values of chromatic dispersion errors				
0.25	DPSK	1.3155 / 0.999988	650	1
	OOK	1.6099 / 0.999976		4
0.5	DPSK	1.2408 / 0.999989		1
	OOK	1.4611 / 0.999980		2
1	DPSK	1.3992 / 0.999986		4
	OOK	1.6073 / 0.999975		4

Table 3.
Summary of the lowest average crosstalk error values and R^2 fit coefficients obtained in the network training process.

Bit delay	Modulation	Average Crosstalk error / R^2	Number of epochs	Serial number of fully connected layer
average values of transfer errors for the highest R^2 coefficients				
0.25	DPSK	0.0858 / 0.999588	250	1
	OOK	0.2651 / 0.986430	250	5
0.5	DPSK	0.0609 / 0.999828	350	4
	OOK	0.2313 / 0.988605	750	5
1	DPSK	0.0688 / 0.999754	750	4
	OOK	0.2341 / 0.988497	450	5
R^2 coefficients for the smallest average values of crosstalk errors				
0.25	DPSK	0.0808 / 0.999537	650	7
	OOK	0.2549 / 0.985635		1
0.5	DPSK	0.0593 / 0.999825		6
	OOK	0.2282 / 0.987843	750	7
1	DPSK	0.0677 / 0.999738		3
	OOK	0.2281 / 0.988102	650	6

Table 4.
Summary of the lowest average OSNR error values and R^2 fit coefficients obtained in the network training process.

Bit delay	Modulation	Average OSNR error / R^2	Number of epochs	Serial number of fully connected layer
average OSNR error values for the highest R^2 coefficients				
0.25	DPSK	0.6607 / 0.984936	650	5
	OOK	0.4832 / 0.989758	650	1
0.5	DPSK	0.5283 / 0.990710	100	2
	OOK	0.3928 / 0.993025	350	5
1	DPSK	0.5434 / 0.989434	450	
	OOK	0.4759 / 0.990934	250	2
R^2 coefficients for the smallest average OSNR error values				
0.25	DPSK	0.6607 / 0.984936	650	5
	OOK	0.4803 / 0.989493	850	1, 5
0.5	DPSK	0.5109 / 0.990507	250	4
	OOK	0.3927 / 0.992879	450	1
1	DPSK	0.5434 / 0.989434		5
	OOK	0.4625 / 0.990773	650	1

The presented models with the best results were further implemented in the method for determining simultaneous impairments.

5. Using learned models to determine simultaneously occurring disturbances

The previous section discussed the trained models. They were used to recognize simultaneously occurring CD, OSNR, and ASE noise phenomena on additional prepared data sets. Each set of 3000 images contained a combination of the mentioned phenomena (CD: 200–1800 ps/nm, crosstalk and ASE noise: 5–40 dB). Tables 5–7 show the accuracy of recognizing the values of phenomena within the permissible range after using the learned models. Values that are underlined indicate the highest accuracy achieved among all sets for each phenomenon. To determine the accuracy of determining value of the individual disturbances, the following (4) was used:

$$accur.(y,\hat{y}) = 100 * \left(\frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} \begin{cases} 1(|\hat{y}_i - y_i| < 0.02 * \hat{y}_i) & \text{for CD} \\ 1(|\hat{y}_i - y_i| \le 0.5) & \text{for Crosstalk, OSNR} \end{cases} \right), \quad (4)$$

where $\hat{y}_i$ is the predicted value of the i-th sample, y_i is the corresponding true value, n_{samples} are all samples in data set.

Some of the values in the tables are expressed in percentages and indicate the number of cases in relation to the entire data set (3000) in which the difference between the reference value and the obtained value was below 2% of the set value for CD and below 0.5 dB for crosstalk and OSNR (requirements for monitoring technicians). For example, the value of 99.63 means that out of 3000 phase portraits tested, the value of the measured phenomenon was correctly determined for 2989 with an acceptable error. Values written as percentages have been rounded to two decimal places, which in some cases may result in a loss of precision, therefore, when converting a percentage value into the number of images, the result should be rounded to the nearest integer. Regarding the set numbers (1 to 7), remember that they represent the number of nodes in the fully connected layers. For example, for set number 1, it is 384/128, i.e., 384 nodes in the first layer and 128 in the second.

The results obtained in the process of determining the values of selected parameters of monitored phenomena using trained CNN models are very good. The flexibility of the ConvNet technique allows for increased accuracy by changing the monitored scope of individual phenomena. As shown in Ref. 26, by reducing the monitoring range for CD from 200–1800 ps/nm to 400–1600 ps/nm and for crosstalk from 5–40 dB to 10–30 dB, it was possible to increase the accuracy of monitoring both phenomena for OOK modulation up to 99.9% for CD and 99.3% for crosstalk. An increase in accuracy was also achieved for ASE noise, and the monitoring accuracy of the OSNR parameter increased to almost 97%.

Table 8 summarizes the results obtained from our own research and those available from Refs. 20 and 21. A direct comparison of the values of individual parameters is not possible because in both available works chromatic and

Table 5.
Comparison of the best results obtained for individual sets in the process of determining the values of parameters describing the monitored phenomena for a delay of 0.25 bit for DPSK and OOK modulation.

	Set number of fully connected layer						
	1	2	3	4	5	6	7
	DPSK modulation, delay 0.25 bit						
CD	99.70 %	99.70 %	**99.83** %	99.73 %	99.73 %	99.73 %	99.60 %
epoch(s)	150	950	150	650	650	750	100, 150, 250, 650
Crosstalk	99.70 %	99.80 %	99.73 %	**99.83** %	99.73 %	99.73 %	99.80 %
epoch(s)	750	650, 750	850	350	950, 1000	150	850
OSNR	94.50 %	94.80 %	94.70 %	94.83 %	**95.03** %	94.8 %	94.83 %
epoch(s)	650, 750	650	750	850	650	650	650
	OOK modulation, delay 0.25 bit						
CD	98.27 %	**98.90** %	98.50 %	98.50 %	98.17 %	98.37 %	98.43 %
epoch(s)	750, 950	250	75	850	850	850	100
Crosstalk	97.30 %	97.03 %	97.13 %	96.90 %	**97.60** %	97.30 %	97.13 %
epoch(s)	850	850	350	250	850	650	150
OSNR	95.40 %	**95.90** %	95.70 %	95.53 %	95.30 %	95.53 %	95.43 %
epoch(s)	750	850	650	950	850	650	750

Table 6.
Comparison of the best results obtained for individual sets in the process of determining the values of parameters describing the monitored phenomena for a delay of 0.5 bit for DPSK and OOK modulation.

	Set number of fully connected layer						
	1	2	3	4	5	6	7
	DPSK modulation, delay 0.5 bit						
CD	99.73 %	99.77 %	99.70 %	**99.80** %	99.50 %	99.77 %	99.77 %
epoch(s)	850	250	250	150	350, 950	250	250
Crosstalk	99.90 %	**99.93** %	99.83 %	99.90 %	99.90 %	99.90 %	99.90 %
epoch(s)	100, 150, 750, 950	150	650, 750	250, 650, 750, 850	650	150, 950	150, 650
OSNR	**95.73** %	95.13 %	**95.73** %	95.40 %	95.10 %	95.60 %	95.50 %
epoch(s)	950	850	750	750, 850	750	650	750
	OOK modulation, delay 0.5 bit						
CD	98.90 %	98.80 %	98.77 %	98.63 %	98.70 %	**98.97** %	98.87 %
epoch(s)	50	650	100	750	750	350	75
Crosstalk	97.50 %	97.30 %	97.50	97.43 %	97.30 %	97.40 %	**97.53** %
epoch(s)	650	250, 750	850	450	250	1000	750
OSNR	96.10 %	96.13 %	96.13 %	96.13 %	96.10 %	**96.40** %	96.33 %
epoch(s)	750	1000	650	650	650, 750	750	750

Table 7.
Comparison of the best results obtained for individual sets in the process of determining the values of parameters describing the monitored phenomena for 1 bit delay and DPSK and OOK modulation.

	Set number of fully connected layer						
	1	2	3	4	5	6	7
	DPSK modulation, delay 1 bit						
CD	99.37 %	**99.47** %	**99.47** %	99.43 %	99.20 %	99.43 %	**99.47** %
epoch(s)	150	150, 650	75	250	950	250	150
Crosstalk	99.80 %	99.83 %	99.73 %	99.80 %	99.80 %	99.80 %	**99.93** %
epoch(s)	250, 650, 1000	250	950, 1000	100	75, 100, 150	750	150
OSNR	95.10 %	94.93 %	95.43 %	**95.50** %	95.30 %	95.30 %	95.10 %
epoch(s)	650, 850	1000	750	750	850	650	950
	OOK modulation, delay 1 bit						
CD	98.47 %	**98.77** %	98.40 %	98.40 %	98.23 %	98.50 %	98.70 %
epoch(s)	350	250	650	150	650	450	100
Crosstalk	97.50 %	97.33 %	97.50 %	**97.60** %	97.40 %	97.40 %	97.20 %
epoch(s)	350, 1000	350	250	950	150, 850	1000	350, 750
OSNR	95.03 %	**95.33** %	95.20 %	94.90 %	94.70 %	94.90 %	95.23 %
epoch(s)	650	650	750	650	650, 850, 950	750	1000

Table 8.
Comparison of the results of our own work with existing results.

Ordinal number	Monitored parameters	Measured range	Average monitoring error	Modulation formats, bitrate
1 source: [20]	OSNR	10 – 28 dB	0.81 dB	QPSK, 16QAM,65QAM 60, 100 Gbit/s
	CD	0 – 450 ps/nm	1.52 ps/nm	
	DGD	0 – 10 ps	0.32 ps	
2 source: [21]	OSNR	10 – 28 dB	0.38 dB	NRZ-DPSK 10 Gbit/s
	CD	0 – 450 ps/nm	1.53 ps/nm	
	DGD	0 – 10 ps	0.48 ps	
	OSNR	10 – 28 dB	0.47 dB	NRZ-OOK 10 Gbit/s
	CD	0 – 450 ps/nm	1.30 ps/nm	
	DGD	0 – 10 ps	0.62 ps	
3 Own results	OSNR	5 – 40 dB	0.53 dB	NRZ-DPSK, 10 Gbit/s
	CD	200 – 1800 ps/nm	1.24 ps/nm	
	Crosstalk	5 – 40 dB	0.06 dB	
	OSNR	5 – 40 dB	0.39 dB	NRZ-OOK, 10 Gbit/s
	CD	200 – 1800 ps/nm	1.46 ps/nm	
	Crosstalk	5 – 40 dB	0.23 dB	

polarization dispersion and ASE noise were monitored simultaneously, while in our own work, CD, crosstalk, and ASE noise were monitored simultaneously. The results presented in Ref. 20 obtained worse values compared to those in Ref. 21, therefore further discussion of these will be omitted. In the results from Ref. 21, for the 10 Gbit/s NRZ-DPSK signal, an OSNR parameter monitoring error of 0.38 dB, CD of 1.53 ps/nm, and DGD of 0.48 ps were obtained. Compared to the results from our own research, the accuracy of the OSNR parameter was improved by 0.15 dB and the accuracy of the CD deteriorated by 0.29 ps/nm. For a 10 Gbit/s NRZ-OOK signal, external testing achieved an OSNR monitoring error of 0.47 dB and 1.30 ps/nm for CD *vs*. OSNR of 0.39 dB and a CD of 1.46 ps/nm for the results from our own work. For this signal, our own research achieved better accuracy for OSNR of 0.08 dB and worse accuracy for a CD of 0.16 ps/nm.

The results presented in Ref. 21 may seem comparable to the results obtained in the conducted research. However, it should be noted that these tests were performed in a much narrower measurement scope of 10–28 dB for OSNR with a step of 2 dB and 0–450 for CD with a step of 50 ps/nm compared to the ranges of 5–40 for OSNR and 200–1800 for CD presented in our own work. In Ref. 21, a small number of phase portraits for the training set of 1100 images was also provided. With the proportions used: 90% – training set and 10% – testing set, the accuracy of the presented method was tested on 110 images. This is not enough. As part of the research, 62 000 images were generated for each case, and the accuracy of the trained models was tested for a test set consisting of 9300 phase portraits (15%). The trained models were then retested using 3000 additional phase portraits. The results presented in Ref. 21 are not very reliable. The experience gained during the research will show that with a data set composed of such a small number of images and a narrow measurement range with a large step between subsequent values, in the event of disturbances other than those in the adopted ranges, the trained model will most likely indicate a very large error. A similar error in the measurement methodology was already observed in Ref. 17, which the authors mentioned in the final conclusions. Additionally, the monitored DGD parameter in the range of 0–10 ps has little impact on the distortions visible within the phase portraits. Polarization dispersion causes two corners of the graph to be slightly rounded. The coexistence of this phenomenon in the presented studies does not have a great impact on the demonstration of the accuracy of the trained models, because mainly the superimposed phenomena of CD and ASE noise are visible there. In the presented own research, the third phenomenon monitored is optical crosstalk in the range of 5 to 40 dB. The demonstrated results obtained in our own research used a very large number of measurement points with a wide range of simultaneously occurring phenomena. An additional check of the trained models on additional data was also applied. This approach ensures that the accuracy of determining parameters describing the monitored phenomena will be very high, regardless of the intensity of other phenomena. This is a very important methodology that has been omitted in the above-mentioned scientific works.

6. Conclusions

As a result of the research, very good results were obtained in determining the values of CD, crosstalk, and ASE noise phenomena. Thanks to the use of the ADTS and CNN methods, it is possible to obtain a high-quality tool for simultaneous monitoring of phenomena occurring in the physical layer of the optical network. Previously published results for determining numerical values using CNN networks in optical monitoring systems were usually completed after a positive learning process [16–22]. It turns out, however, that the highest obtained accuracy of the trained model does not always translate into equally good results in the further process of determining the values and the trained model may monitor parameters with a large error. Therefore, to further assess the quality of the monitoring tool, the trained models should be further validated on additional external data. The presented results complement the missing knowledge regarding monitoring and address the above-mentioned problems. However, they do not exhaust further possibilities for developing and improving the technique. It is advisable to conduct further research towards simultaneous monitoring of a larger number of parameters and optimization of the minimum number of images from datasets necessary for the network training process to maintain high quality monitoring. It is also necessary to complete research in real conditions. The proposed measurement method using convolutional networks required many training data (even tens of thousands) containing a combination of three occurring phenomena with dense coverage. So many phase portraits that would simultaneously contain phenomena of different intensities would be difficult to obtain in real conditions (as opposed to a simulation environment). This would require the construction of a special station to generate data and a very large amount of time to obtain it. For this reason, it was decided to use a simulation environment. After obtaining the very good results presented in this article, it is justified to undertake work that will use real data.

References

[1] Zhao, J. *et al*. NRZ-DPSK and RZ-DPSK signals signed chromatic dispersion monitoring using asynchronous delay-tap sampling. *J. Light. Technol.* **27**, 5295–5301 (2009) . https://doi.org/10.1109/JLT.2009.2031610

[2] Dods, S, D. & Anderson, T. B. Optical Performance Monitoring Technique Using Delay Tap Asynchronous Waveform Sampling. in *2006 Optical Fiber Communication Conference and the National Fiber Optic Engineers Conference* 3 (IEEE, 2006). https://doi.org/10.1109/ofc.2006.215890

[3] Anderson, T. *et al.* Multi impairment monitoring for optical networks. *J. Light. Technol.* **27**, 3729–3736 (2009). https://doi.org/10.1109/JLT.2009.2025052

[4] VPIphotonics: Simulation Software and Design Services https://www.vpiphotonics.com/index.php (accessed Feb. 8, 2024).

[5] Chan, C. *Optical Performance Monitoring*. (Elsevier Inc., 2010).

[6] Anderson, T. *et al.* Demonstration of Simultaneous OSNR and CD Monitoring Using Asynchronous Delay-Tap Sampling on an 800 km WDM Test Bed. in *35th European Conference on Optical Communication (ECOC 2009)* 2–3 (VDM Verlag, 2009).

[7] Choi, H. Y., Takushima, Y. & Chung, Y. C. Multiple-Impairment Monitoring Technique Using Optical Field Detection and Asynchronous Delay-Tap Sampling Method. in *Conference on Optical Fiber Communication, Technical Digest Series* OThJ5 (OSA, 2009).

[8] Zhao, J. *et al*. A Novel Optical Signal Monitoring Method of DPSK Signal Based on Delay Tap Sampling and Hausdorff Distance Measure. in *2008 Conference on Quantum Electronics and Laser Science Conference on Lasers and Electro-Optics, CLEO/QELS* 1–2 (OSA, 2008). https://doi.org/10.1109/CLEO.2008.4552261

[9] Jargon, J. A, Wu, X. & Willner, A. E. Optical Performance Monitoring by Use of Artificial Neural Networks Trained with Parameters Derived From Delay-Tap Asynchronous Sampling. in *Optical Fiber Communication Conference 2009* 3–5 (OSA, 2009). https://doi.org/10.1364/ofc.2009.othh1

[10] Kozicki, B., Takuya, O. & Hidehiko, T. Optical performance monitoring of phase-modulated signals using asynchronous amplitude histogram analysis. *J. Light. Technol.* **26**, 1353–1361 (2008). https://doi.org/10.1109/JLT.2008.917374

[11] Kozicki, B., Maruta, A. & Kitayama, K. Experimental demonstration of optical performance monitoring for RZ-DPSK signals using delay-tap sampling method. *Opt. Express* **16**, 3566–3576 (2008). https://doi.org/10.1364/oe.16.003566

[12] Kozicki, B., Maruta, A. & Kitayama, K. Transparent performance monitoring of RZ-DQPSK systems employing delay-tap sampling. *J. Opt. Netw.* **6**, 1257–1269 (2007). https://doi.org/10.1364/JON.6.001257

[13] LeCun, Y., Bottou,. Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2323 (1998). https://doi.org/10.1109/5.726791

[14] Chen, W. K. *The Electrical Engineering Handbook*. (Elsevier Academic Press, 2005).

[15] LeCun, Y., Kavukcuoglu, K. & Farabet, C. Convolutional Networks and Applications in Vision. in *ISCAS 2010–2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems* 253–256 (IEEE, 2010). https://doi.org/10.1109/ISCAS.2010.5537907

[16] Zhang, J. *et al.* Intelligent adaptive coherent optical receiver based on convolutional neural network and clustering algorithm. *Opt. Express* **26**, 18684–18698 (2018). https://doi.org/10.1364/oe.26.018684

[17] Cheng, Y., Fu, S., Tang, M. & Liu, D. Multi-task deep neural network (MT-DNN) enabled optical performance monitoring from directly detected PDM-QAM signals. *Opt. Express* **27**, 19062–19074 (2019). https://doi.org/10.1364/oe.27.019062

[18] Wang, D. *et al.* Intelligent constellation diagram analyzer using convolutional neural network-based deep learning. *Opt. Express* **25**, 17150–17166 (2017). https://doi.org/10.1364/oe.25.017150

[19] Wang, D. *et al.* Modulation format recognition and OSNR estimation using CNN-based deep learning. *IEEE Photon. Technol. Lett.* **29**, 1667–1670 (2017). https://doi.org/10.1109/LPT.2017.2742553

[20] Fan, X. *et al.* Feature fusion-based multi-task convnet for simultaneous optical performance monitoring and bit-rate/modulation format identification. *IEEE Access* **7**, 126709–126719 (2019). https://doi.org/10.1109/ACCESS.2019.2939043

[21] Fan, X. *et al.* Joint optical performance monitoring and modulation format/bit-rate identification by CNN-based multi-task learning. *IEEE Photon. J.* **10**, 1–12 (2018). https://doi.org/10.1109/JPHOT.2018.2869972

[22] Wan, Z. *et al*. Intelligent optical performance monitor using multi-task learning based artificial neural network. *Opt. Express* **27**, 11281–11291 (2019). https://doi.org/10.1364/OE.27.011281

[23] TensorFlow White Papers https://www.tensorflow.org/about/bib (accessed Oct. 27, 2020). (in Polish)

[24] Aymeric, D. *et al.* TFLearn. *GitHub* https://github.com/tflearn/tflearn (2016). (accessed Jan. 25, 2022).

[25] Skarbek, W. Symbolic tensor neural networks for digital media : from tensor processing via BNF graph rules to CREAMS applications. *Fundam. Inform.* **168**, 89–184 (2019). https://doi.org/10.3233/FI-2019-1827

[26] Mrozek, T. & Perlicki, K. Simultaneous monitoring of the values of CD, crosstalk and OSNR phenomena in the physical layer of the optical network using CNN. *Opt. Quant. Electron.* **53**, 640 (2021). https://doi.org/10.1007/S11082-021-03280-5