IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 22 November 2022, accepted 11 December 2022, date of publication 19 December 2022,
date of current version 28 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3230834

RESEARCH ARTICLE

Joint Modulation Format Identification and Optical Signal-to-Noise Ratio Monitoring Based on Ternary Neural Networks

PENG ZHOU[1], CHUANQI LI[1,2], DONG CHEN[1,2], YU ZHANG[3], AND YE LU[3]
[1]College of Computer Science & Information Engineering, Guangxi Normal University, Guilin 541000, China
[2]College of Physics and Electronic Engineering, Nanning Normal University, Nanning 530000, China
[3]College of Electronic Engineering, Guangxi Normal University, Guilin 541000, China

Corresponding author: Ye Lu (luye@mailbox.gxnu.edu.cn)

This work was supported by Guangxi S&T Program under Grant AB17292082.

ABSTRACT Modulation format identification (MFI) and optical signal-to-noise ratio (OSNR) monitoring are essential for elastic optical networks. A method for joint modulation format identification and OSNR monitoring based on ternary neural networks was proposed. Further, a ternary neural network was established, and the constellation images after constant modulus algorithm (CMA) equalization were used as input features. Four commonly used modulation formats were distinguished, including dual-polarization (DP)-QPSK, 8QAM, 16QAM, and 64QAM. A 32G baud simulation system was constructed, and the results were analyzed. First, we investigated the influence of resolution and sample length on the identification accuracy. The values 32 and 7000 were selected for the two parameters, respectively, based on the balance between resource consumption and accuracy. Then, we compared the performance of the binary neural network (BNN), ternary neural network (TNN), and full-precision neural network (FNN). The results indicate that the memory consumption and extraction time of the TNN are similar to those of the BNN, and the accuracy is further improved. Moreover, the robustness of the system was analyzed. The results validate that the proposed method can tolerate fiber nonlinearity (from 500 km to 1500 km). Finally, an experiment was conducted to prove the practicality of the proposed method.

INDEX TERMS Modulation format identification (MFI), optical signal-to-noise ratio (OSNR) monitoring, ternary neural network (TNN), optical fiber communication.

I. INTRODUCTION

With the iteration of the Internet of Things technology, lifestyle has evolved from traditional to modern; this is evident in innovations, such as autonomous vehicles and the wise information technology of med (WITMED) [1], [2], [3]. Optical networks are required for high spectrum utilization, flexible structure, specific quality of service (QoS), and accurate optical performance monitoring (OPM) [4], [5].

During the long haul transmission, signals suffer various degradation, such as chromatic dispersion (CD), polarization mode dispersion (PMD), frequency offset, and polarization dependent-loss (PDL) [6], [7]. Recently, digital signal processing (DSP) technology has been adopted in the coherent receiver to compensate for the impact of impairments after coherent detection. In the DSP module, there are many algorithms, including CD equalization algorithm, time recovery algorithm, constant modulus algorithm (CMA), multi-modulus algorithm (MMA), frequency offset compensation algorithm, carrier phase recovery algorithm, and decoder. All algorithms can be divided into the following two categories: modulation format dependent algorithms and modulation format independent algorithms. Modulation format-independent algorithms (i.e., CD equalization, time recovery, and CMA) are applicable to signals of all MFs, whereas modulation

The associate editor coordinating the review of this manuscript and approving it for publication was Jiafeng Xie.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

format dependent algorithms (i.e., MMA, frequency offset compensation, carrier phase recovery, and decoder) require MFs as the prior information, and the corresponding algorithm is selected for signal processing [8]. Hence, an accurate modulation format identification (MFI) algorithm is directly related to the validity of the demodulation. In addition, the optical signal-to-noise ratio (OSNR) is considered to be a key parameter of optical communications [9], which has a direct relationship with bit error rate (BER). Through monitoring the OSNR, we can maintain track of the channel and take measures for the degraded nodes. Therefore, efficient MFI and OSNR estimation algorithms are essential for optical communications.

In recent years, artificial intelligence (AI) has been widely used in various fields, including optical communications. Various AI technologies are applied in MFI. The look-up table and neural networks are used to distinguish MFs [10]. The method is based on the amplitude variance. In [3], [11], [12], and [13], an MFI method based on the artificial neural network utilizes amplitude histograms (AHs) as the recognition feature. In [14], a CNN-based cost-effective MFI method is proposed. The highlight of the scheme is to obtain features through low bandwidth detecting and low rate sampling and thus reduce the implementation cost significantly. To improve the functionality of the system, various neural networks are used to identify MFs and monitor OSNR simultaneously. In [15], [16], and [17], random forest, Gaussian process regression, support vector machine, and neural network, etc., were used for modulation format recognition and transmission quality monitoring. Zhang et al. proposed a cascaded neural network based on transfer learning to simultaneously identify the modulation formats and monitor the OSNR values [18]. They completed the MFI based on amplitude histograms (AHs) in the first-level neural network. Then, the MFI and AHs were used as input features to accelerate the training of the second-level neural network. However, using AHs as input features increases the complexity of the algorithm. In addition, they did not analyze the robustness (e.g., CD and nonlinearity) of the model. Ahmed K. Ali et al. proposed a unique modulation classification method to improve the extracted features employed for the recognition of modulation schemes [19]. The proposed method combines higher-order cumulants with threshold classifiers, and improves the cumulant features using the properties of a natural logarithmic function, which make the classifier efficient with respect to decision-making. To shorten training time and reduce running memory, Zhao et al. proposed a method of OPM and MFI based on the binary neural network (BNN) [20]. In the scheme, the accuracies of MFI and OSNR estimations are 100% and 97.71%, respectively, which are close to those of the float-valued CNN and multi-layer perceptron. BNN significantly reduces memory consumption and saves training time. However, the OSNR accuracy decreases by 9% when the launched power is 7 dBm because the BNN compresses the model significantly.

In this study, we propose a method based on a ternary neural network (TNN), which can simultaneously identify modulation formats and monitor OSNR. In the proposed scheme, constellation images processed by the CMA algorithm are chosen as input features and the TNN is trained. To verify the performance of our algorithm, we first demonstrate a 32G baud dual-polarization (DP) coherent transmission system based on commercial software (VPI transmission maker) with four commonly used MFs (i.e., QPSK, 8QAM, 16QAM, and 64QAM). Then, we investigated the influence of resolution and sample length on the identification accuracy. The values 7000 and 32 were selected for the two parameters, respectively, by considering the resource consumption and accuracy. Furthermore, we compared the performance of the binarized neural network (BNN), ternary neural network (TNN), and full-precision neural network (FNN). The TNN can achieve a balance between resource consumption and accuracy. The results show that the TNN can achieve an accuracy similar to that of the FNN and better than that of BNN. However, compared with the FNN, the TNN saves running memory and reduces storage and computing power consumption. In addition, the robustness of the system was analyzed, and the results validate that the proposed scheme can tolerate fiber nonlinearity in a specified range. Finally, an experiment is conducted to prove the practicality of the proposed scheme.

II. OPERATION PRINCIPLE

A. TENERAL NEURAL NETWORKS

TNN outperforms BNN and FNN in terms of expressive ability, model compression, and computational requirement. The TNN compresses the weight into three values, which require 16× to 32× less memory usage than the FNN and 2× more memory usage than the BNN. However, the TNN has gained 38× more expressive abilities than the BNN in most recent network architectures [21]. Moreover, the weights of the TNN add an extra 0 state, and multiplication is replaced by addition during the computation. Therefore, the 0 state does not need to be calculated during the addition process, and the occupied computational resources are almost similar to those of BNN [22]. In conclusion, the TNN is theoretically more practical than the BNN.

In ternary neural networks, the ultimate goal is to replace full precision floating weights W with ternary weights Wt, while reducing the Euclidean distance between W and Wt along with a non-negative scaling factor α. The optimization problem is expressed in Eq. 1 below.

$$\begin{cases} \alpha^*, W^{t*} = \arg\min\limits_{\alpha W^t}, & J(\alpha, W^t) = \left\| W - \alpha W^t \right\|_2^2 \\ s.t. \alpha \geq 0, & W_i^t \in \{-1, 0, +1\}, i = 1, 2, \ldots, n \end{cases} \tag{1}$$

One way to reduce the loss of the TNN is to introduce zero as a quantized value. Thus, we selected a threshold value to constrain the full-precision weights to -1, 0, and +1.

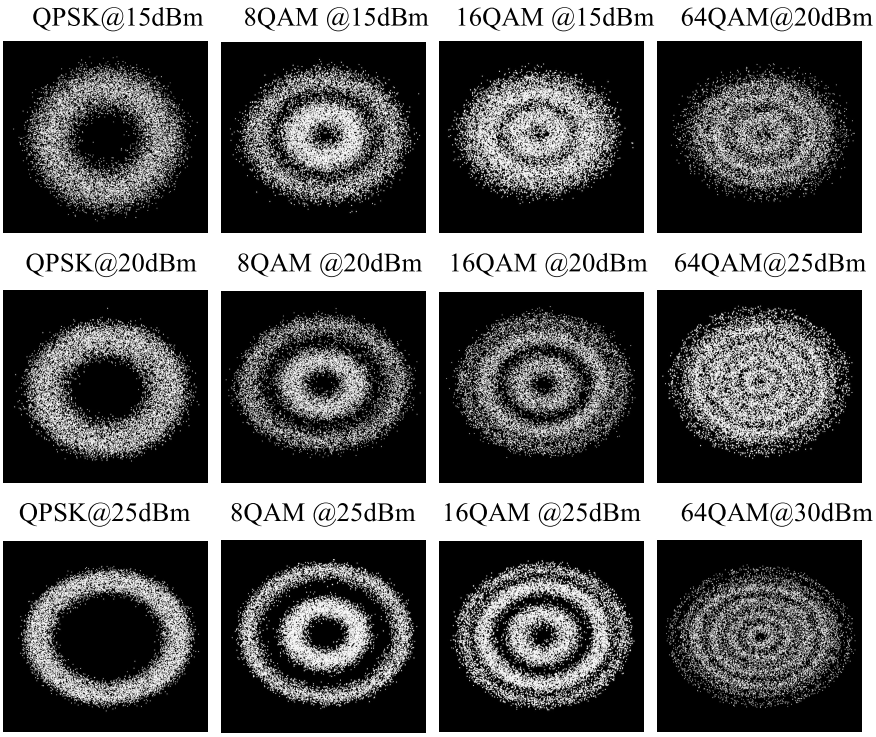

FIGURE 1. Gray images of constellations with different MFs and OSNR.

The principle is given as follows:

$$W_i^t = f_t(W_i|\Delta)\begin{cases} +1, & if\ W_i > \Delta \\ 0, & if\ |W_i| \le \Delta \\ -1, & if\ W_i < -\Delta \end{cases} \tag{2}$$

where Δ is a positive threshold value. W_i is the full precision floating weight of layer i and W_i^t represents the ternary weight of layer i. By combining (1) and (2) for any given Δ, optimization α_Δ^* can be converted as follows:

$$\alpha_\Delta^* = \frac{I}{|I_\Delta|}\sum_{i\in I_\Delta}|W_i| \tag{3}$$

where $I_\Delta = \{i|\,|W_i| > \Delta\}$ and $|I_\Delta|$ is the number of elements in I_Δ. In the proposed method, Δ is expressed as

$$\Delta* = 0.7 * E(W) \approx \frac{0.7}{n}\sum_{i=1}^{n}|W_i|. \tag{4}$$

The forward propagation of the ternary neural network is given in Eq. 5.

$$\begin{cases} Z = X * W \approx X * (\alpha W^t) = (\alpha X) \oplus W^t \\ X^{next} = g(Z) \end{cases} \tag{5}$$

where X and X^{next} represent the input and output of the block, respectively. $g()$ is a nonlinear activation function. XOR can be implemented by addition. Specifically, the activated output of the previous block consists of ternary values including 0, and ± 1, which will be added with the ternary weight of next layer. Therefore, the calculation of ternary values simplifies the forward and backward propagation by turning multiplication into addition.

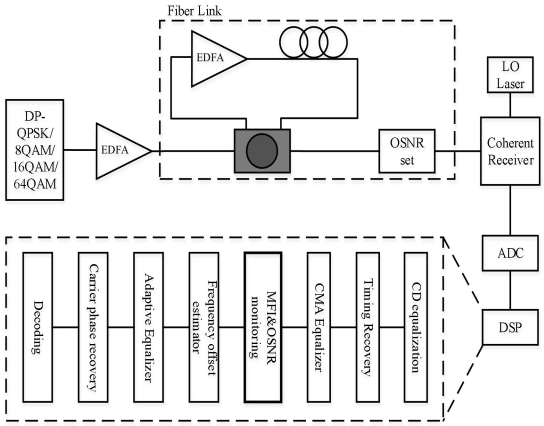

FIGURE 2. Principle of the optical communication system with the proposed method.

To evaluate the success of the proposed models, predictive value represented by precision was studied [23]. The performance index metrics can be defined as follows:

$$Precision(\%) = \frac{TP}{TP + FP} \tag{6}$$

where TP indicate the ratio of signals that are correctly classified, and FP indicate the ratio of signals that are incorrectly classified.

TABLE 1. Label vectors of MFs and OSNR values.

Signal type	Label vector
QPSK, 10 dBm	+1,0,0,0 $\underbrace{+1,0,\ldots,0}_{16}$
QPSK, 11 dBm	+1,0,0,0 $\underbrace{0,+1,\ldots,0}_{16}$
QPSK, 25 dBm	+1,0,0,0 $\underbrace{0,0,\ldots,+1}_{16}$
8QAM, 10 dBm	0,+1,0,0 $\underbrace{+1,0,\ldots,0}_{16}$
16QAM, 29 dBm	0,0,+1,0 $\underbrace{0,\ldots,+1,0}_{16}$
64QAM,35 dBm	0,0,0,+1 $\underbrace{0,0,\ldots,+1}_{16}$

B. IDENTIFICATION PRINCIPLE

In the fixed optical network, the flexible transmitter is suitable for the multi QoS. Different modulation formats are allocated according to different requirements to improve spectrum efficiency. In our method, commonly used formats, such as QPSK, 8QAM, 16QAM, and 64QAM were included. In the electric field, the ideal QPSK, 16QAM, and 64QAM symbols can be expressed as [24]

$$\begin{aligned} E_{QPSK}(k) &= A\exp(i\varphi_k), \varphi_k = \frac{(2m-1)\pi}{4}, m = 1,2,3,4 \\ E_{16QAM}(k) &= a_k + ib_k, \quad a_k, b_k = \pm 1, \pm 3 \\ E_{64QAM}(k) &= a_k + ib_k, \quad a_k, b_k = \pm 1, \pm 3, \pm 5, \pm 7 \end{aligned} \tag{7}$$

where the amplitudes of the MFs are uniformly distributed. The grayscale images of the constellation with different modulation formats and OSNR after the CMA equalization are shown in Fig. 1.

Different MFs have different amplitude distribution characteristics. There are obvious amplitude levels. The levels for QPSK, 8QAM, 16QAM, and 64QAM are 1, 2, 3, and 9, respectively. Moreover, the amplitude distribution characteristics become more obvious when the OSNR increases. Thus, the grayscale images of the constellation were used as the input features for the neural network training in the proposed scheme.

III. SIMULATION, RESULTS, AND DISCUSSIONS

A. SIMULATION SCHEME AND DATA PREPRATION

To verify the correctness and reliability of the algorithm, we designed an optical communication system based on VPI and PyCharm as commercial software. Fig. 2 shows the principle of the optical communication system with the proposed algorithm. At the transmitter, 32GBaud dual-polarization (DP)-QPSK, 8QAM, 16QAM, or 64QAM signals were generated. The wavelength of the laser was 1550 nm with a line width of 100 KHz, and the range of launch power was 0–5 dBm. The fiber link of the system comprised an erbium-doped fiber amplifier (EDFA), OSNR set, and standard single-mode fiber (SSMF). The noise figure of the EDFA is 3.8dBm. The length of fiber was 100 km per span, and the CD, PDM, non-linear, and attenuation coefficients were set as 16 ps/(km.nm), 0.1 ps/km^{-2}, 2.6×10^{-20} s/m^3, and 0.2 dBm/km, respectively. The OSNR set block was set in the range of 10–35 dBm. After the long haul transmission, the signal was mixed with the local oscillator and compensated in the DSP module after the analog to digital conversion. In the DSP, the CD equalization, timing recovery, CMA equalization, MFI and OSNR monitoring, MMA equalization, frequency offset estimator, adaptive equalizer, carrier phase recovery, and decoding were sequentially executed.

Based on the above principles, 10 groups of 2×10^5 symbols were generated at the transmitter. In each group of symbols, 10 groups of data with the length of 15000 were selected randomly. Then, the 15000 data symbols were used to generate grayscale images with the same OSNR and MF. In the proposed method, each MF contains 16 OSNR values (10–25 dBm for DP-QPSK and DP-8QAM, 15–30 dBm for DP-16QAM, and 20–35 dBm for DP-64QAM). Hence, the data set contains $100 \times 16 \times 4 = 6400$ grayscale images. Next, we randomly selected 70% as the training set, 20% as the verification set, and 10% as the test set. The output of the neural network was a 20 bits sequence consisting of +1 and 0. As shown in Table 1, +1 represents the label vector of the MF (in the first 4 bits) and OSNR (in the last 16 bits).

B. STRUCTURE OF TENERAL NEURAL NETWORKS

We build up the ternary neural network model like the structure in [25], which contains three convolutional layers, one average pooling layer, and one fully-connected layer. Fig. 3 shows the structure of the designed ternary neural network. The three convolutional layers have 64, 64, and 128 filters (convolution kernel), respectively. The size of all the filters was 3×3. Then, an average pooling layer was followed to reduce the parameter size. The one fully-connected layers contained 20 nodes, respectively. The loss function adopts the cross-entropy loss function which is widely used in classification tasks. The loss function for a single task is as follows:

$$Loss = -\frac{1}{\mathrm{n}}\sum_{i=1}^{\mathrm{n}} Y_i \log y_i \tag{8}$$

where n is the number of neurons. Y_i is the true probability distribution and Y_i is the prediction probability distribution. The loss function for two tasks is given by Eq. 8 below.

$$\begin{aligned} l = \lambda_1 l_1 + \lambda_2 l_2 = -\frac{1}{\mathrm{n}}\lambda_1 \sum_{i=1}^{n} Y_i^{\mathrm{m}} \log y_i^m \\ -\frac{1}{\mathrm{N\text{-}n}}\lambda_2 \sum_{i=1}^{N-n} Y_i^o \log y_i^o \end{aligned} \tag{9}$$

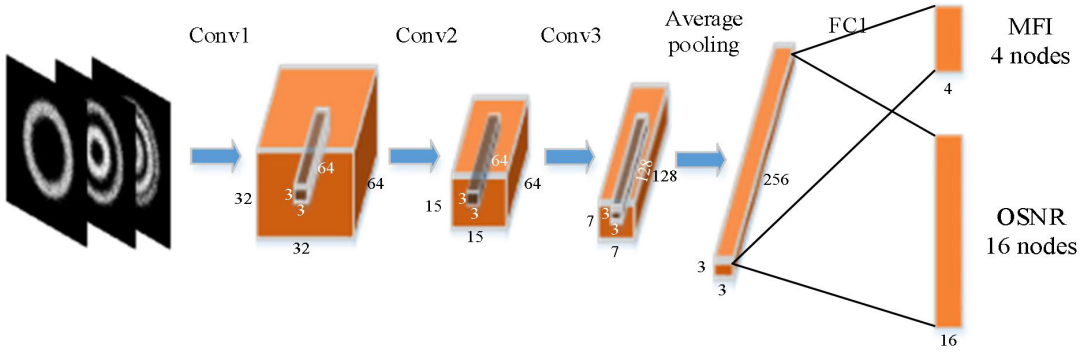

FIGURE 3. Structure of the ternary neural network.

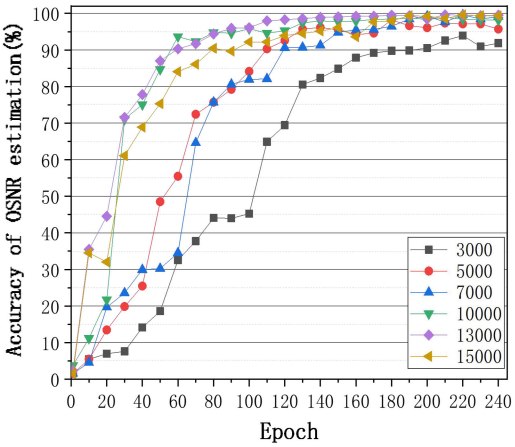

FIGURE 4. Accuracy of OSNR estimation varying with sample lengths.

where λ_1 and λ_2 are the task weights of the MFI and OPM, respectively. N represents the nodes of the output layer, and n is nodes for MFI. Y_i^m is the true probability distribution for the MFI and y_i^m is the prediction probability distribution for the MFI. Y_i^o is the true probability distribution for the OPM and y_i^o is the prediction probability distribution for the OPM.

As a comparison, BNN and FNN were built with similar architecture and loss function as TNN.

C. INFLUENCE OF SAMPLE LENGTH AND RESOLUTION

In this section, we investigated how the sample length and the resolution of grayscale map influence the identification accuracy. We fixed the fiber length at 200km, and launch power at 0dBm. The grayscale images are generated according to section III.

The number of elements in each grayscale image has a significant impact on the recognition results. We studied the effect of sample length (in the range of 3000 to 15000) on system performance. Here, we denote the sample length by L.

Fig. 4 illustrates how the OSNR accuracy varies with sample length. It is evident that the OSNR accuracy increases gradually with an increase in epoch. However, the curves of different sample length rise at different rates. The performance of $L \geq 7000$ is similar when the epoch is equal to 160. The OSNR accuracy of $L < 7000$ is lower than $L \geq 7000$ when the curves are flat. Fig. 5 shows the loss values of different sample lengths at different epochs. When the epoch is equal to 220, there is a little difference between the loss values of 7000, 10000, 13000, and 15000. Hence, we set $L = 7000$ after considering the accuracy and complexity of the algorithm.

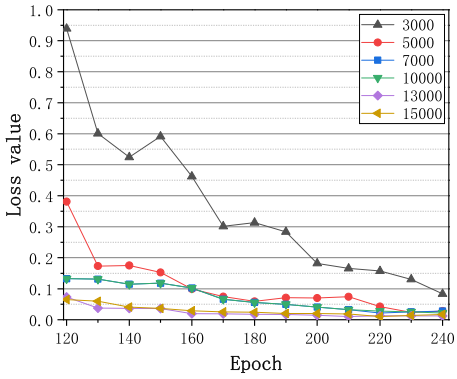

FIGURE 5. Loss values of different sample lengths at different epochs.

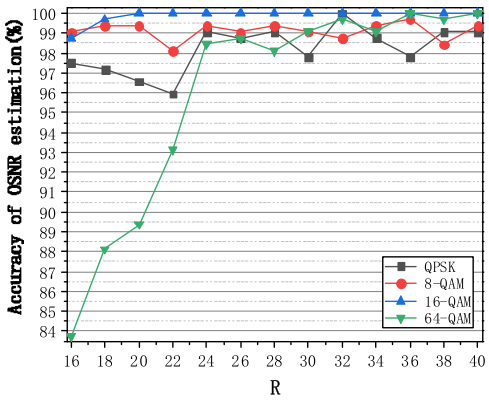

FIGURE 6. OSNR accuracy of different MFs with different R.

Then, we examined the influence of resolution which is the number of pixels in each image. We assumed that

the resolution is $R \times R$. To extract the features from the constellation, the images used as input features must retain more details when input into the TNN. When the resolution is too small, some important details become mosaics. This greatly affects the accuracy of the MFI and OSNR estimation. We set R in the range of 16 to 40. Fig. 6 shows the OSNR accuracy of different MFs with varying R in the range of 16 to 40. It is evident that when resolution is equal to 32, the OSNR accuracy of QPSK, 16QAM, and 64QAM is close to 100%, and the OSNR accuracy of 8QAM is close to 99%, which meets the requirements of an optical communication system. Thus, $R = 32$ is selected as the resolution to achieve a balance between performance and computation.

TABLE 2. Performance comparison of different CNNs.

Model	FNN	BNN	TNN
Memory size(MB)	3.62	0.14	0.19
Execution time(s)	1.91	0.77	0.89

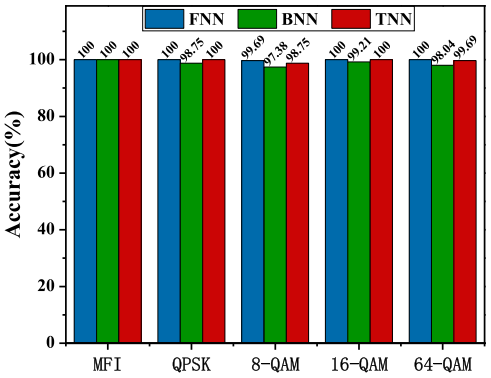

FIGURE 7. Accuracy of MFI and OSNR estimation for FNN, BNN, and TNN.

D. COMPARISION OF DIFFERENT CNNs

Table 2 shows the performance comparison of different CNNs when {L,R}={7000,32}. Two important aspects are compared, including the memory size and execution time. Compared with FNN, BNN and TNN significantly reduce the memory size and execution time, which is benefited from the weight compression of the BNN and TNN. Fig. 7 shows the accuracy of the MFI and OSNR estimations for FNN, BNN and TNN. It is expected that the MFI accuracy of the three models are 100%. For OSNR estimation, the accuracy of the TNN is higher than that of the BNN, and is closer than FNN. Therefore, we can conclude that the TNN is more suitable for MFI and OPM tasks. Fig.8 shows the real OSNR values vs predicted OSNR results. It can be observed that the false predicted values for 8QAM are concentrated in the range of 15 dBm to 17 dBm, and the error is in the range of 2dBm. There is a large gap between the false predicted values and the true values for 64QAM and it may be an accidental error. In general, our scheme can achieve good performance in MFI and OSNR estimation.

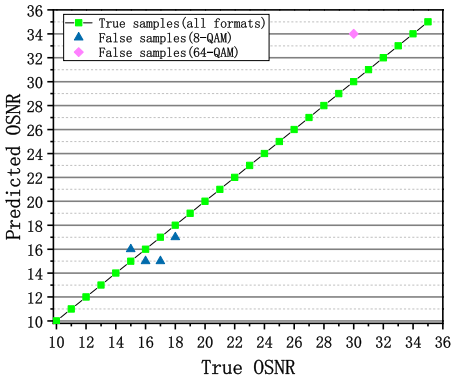

FIGURE 8. The real OSNR values vs predicted OSNR results.

Table 3 shows the comparison of several recent methods. It is clear that the MFI accuracy of each method is 100% and the OSNR accuracy of each method is close to 100%. However, the method presented in this paper uses the minimum sample length and total parameters.

E. ROBUSTNESS EVALUATION

Fiber nonlinearity is an important factor in long-haul optical communication systems. The nonlinear effect has an intimate relationship with the transmission distance and launch power. The launch power was set as 3 dBm, and the fiber length was set in the range of 500 to 1500 km at a 200 km interval. 10 constellation diagrams for each distance were generated.

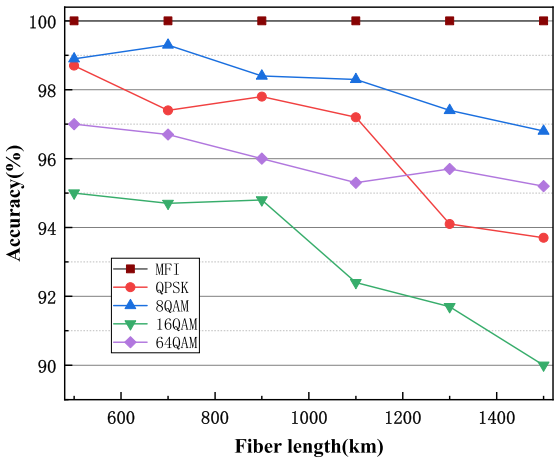

FIGURE 9. Accuracy of MFI and OSNR estimation at the different fiber lengths.

Fig. 9 shows the accuracy of the MFI and OSNR estimations at different fiber lengths. From the diagram, it can be observed that the accuracy of the MFI is 100%, and the accuracy of the OSNR estimation for QPSK, 8QAM, 16QAM, and 64QAM is is above 90%. When the distance increases, the OSNR accuracy of each modulation format decreases by 4 percentage points. This confirms that the proposed method has a certain tolerance for nonlinearity.

TABLE 3. Comparison with other systems in the literature.

References	Modulation formats	OSNR range	Data rate /Gbaud	Model	Sample length	MFI accuracy	OSNR estimation	Total parameter
[18]	DP-QPSK	10 - 25	12.5	TL-CSNN	32768	100%	E_{RMSE} = 0.29 dB	119627
	DP-6QAM							
	DP-16QAM	15 - 30						
	DP-32QAM	20 - 36						
	DP-48QAM							
	DP-64QAM							
[20]	DP-QPSK	10 - 25	12.5	BNN	2×10^5	100%	99.17%	153396
	DP-6QAM						97.71%	
	DP-8QAM						99.65%	
	DP-12QAM						98.82%	
	DP-16QAM	15 - 30					98.19%	
	DP-24QAM						98.33%	
	DP-32QAM	20 - 35					99.44%	
	DP-48QAM						99.17%	
	DP-64QAM						99.72%	
[25]	DP-QPSK	10 - 25	10	CNN	1×10^4	100%	100%	411732
	DP-8QAM							
	DP-16QAM	15 - 30						
	DP-64QAM	20 - 35						
[26]	QPSK	10 - 26	0.155	CNN	65536	100%	97.6%	188307
	8QAM						100%	
	16QAM						100%	
Proposed	DP-QPSK	10 - 25	32	TNN	7000	100%	100%	116564
	DP-8QAM						98.75%	
	DP-16QAM	15 - 30					100%	
	DP-64QAM	20 - 35					99.69%	

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

To verify the feasibility of the proposed method, a proof-of-concept system was constructed. Fig. 10 shows the principle of the dual-polarization transmission system. The external cavity laser emitted light with a center wavelength of 1549.5 nm and a line width of 100 kHz. The arbitrary waveform generator was used to drive the I/Q modulator. 10G baud DP-QPSK/8QAM/16QAM/64QAM signals were generated at the transmitter. The EDFA1 maintained the output power at 0 dBm. The link includes N spans 100 km SMF and EDFA. EDFA2 and variable optical attenuator varied the OSNR of the signals in the range of 10 to 36. At the receiver, the 90° optical hybrid performed coherent mixing of the local oscillator light and signal light. Then, the signal was sampled using a digital oscilloscope with a sampling rate of 50 GSa/s after photoelectric detection. Finally, the digital signal was processed using offline DSP technology.

First, we set the transmission distance at 100–500 km with an interval of 100 km.The launch power was set as 3 dBm. Similar to the method used in Section III, the gray images of the constellation were generated based on the data from different distances. For each distance, 6400 images were collected, and the data set contained 32000 images. Fig. 11 shows the accuracy of MFI and OSNR estimations obtained by the TNN trained with data from different distances. For the MFI, the proposed method can reach 100% accuracy when the distance parameters are dynamic, which implies that the MFI maintains robustness. For the OSNR

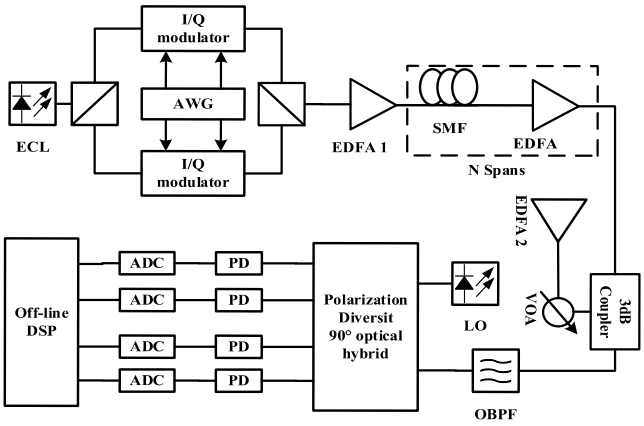

FIGURE 10. Principle of the proof-of-concept system.

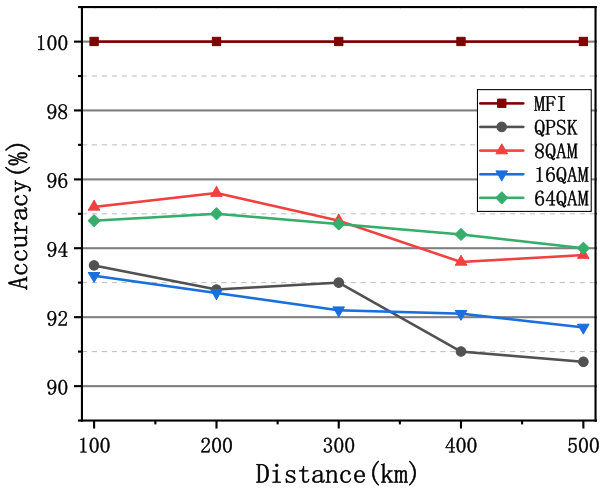

FIGURE 11. Accuracy of MFI and OSNR estimation by TNN trained with data from different distances.

estimation, the accuracy of the four MFs decreases to a certain extent.

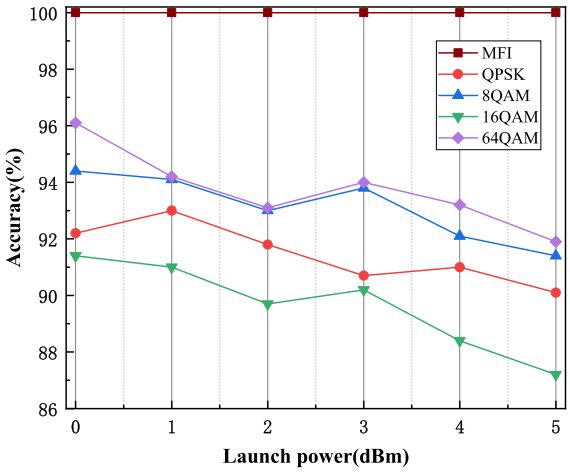

FIGURE 12. Accuracy of MFI and OSNR estimation by TNN trained with data from different launch power.

Then, we set the launch power at 0–5 dBm with an interval of 1 dBm, and the fiber length was 500 km. The gray images of the constellation were generated based on the data from different launch power. For each launch power, 6400 images were collected, and the data set contained 32000 images. Fig. 12 shows the accuracy of the MFI and OSNR estimations by TNN trained with data from different launch power. Similar to the case of different distances, the MFI attained 100% accuracy, which implies that our method maintains robustness for different launch power in the MFI task. However, the results of the OSNR estimation task are not satisfactory. The accuracy of each modulation format dropped by 6-7%. From the above research, it is necessary to conduct further research by exploring a more practical CNN to achieve a balance between performance and complexity and developing a parameter-insensitive CNN that does not require repeated training when the parameters are changed.

V. CONCLUSION

In this study, a method based on TNN that simultaneously identifies modulation formats and monitors OSNR was proposed. Four commonly used MFs (including DP-QPSK/8QAM/16QAM and 64QAM) and 16 OSNR values for each MFs were investigated. Constellation images after the CMA equalization were used as input features for the training of the TNN. A 32G baud dual-polarization coherent transmission system was demonstrated and the influence of resolution and sample length on the identification accuracy was investigated. The values 32 and 7000 were selected for the two parameters, respectively. Then, three models (FNN, BNN, and TNN) are compared in terms of memory size, execution time, and accuracy. The results show that the TNN can achieve better results with fewer resources. Moreover, we investigated the effects of nonlinearity on the system, and the results prove that our method exhibits better robustness. Finally, a proof-of-concept system was constructed to ensure the practicality of our method. There are several aspects need to be further explored. Firstly, compared with BNN, the TNN has improved the accuracy of OSNR estimation. However, it has not reached 100% and the further improvement should be taken in practical application. Secondly, the memory size and execution time of TNN is greater than those of BNN. It is worth doing further study (such as introducing dilated convolution) to solve this problems. Although this scheme requires further study, it can a provide reference for future research.

REFERENCES

[1] A. Yi, H. Liu, L. Yan, L. Jiang, Y. Pan, and B. Luo, "Amplitude variance and 4th power transformation based modulation format identification for digital coherent receiver," *Opt. Commun.*, vol. 452, pp. 109–115, Dec. 2019.

[2] R. Zhao, W. Sun, H. Xu, C. Bai, X. Tang, Z. Wang, L. Yang, L. Cao, Y. Bi, X. Yu, and W. Fang, "Blind modulation format identification based on improved PSO clustering in a 2D Stokes plane," *Appl. Opt.*, vol. 60, no. 31, pp. 9933–9942, 2021.

[3] S. Li, J. Zhou, Z. Huang, and X. Sun, "Modulation format identification based on an improved RBF neural network trained with asynchronous amplitude histogram," *IEEE Access*, vol. 8, pp. 59524–59532, 2020.

[4] L. Yang, H. Xu, C. Bai, X. Yu, K. You, W. Sun, X. Zhang, and C. Liu, "Low-complexity modulation format identification scheme via graph-theory in digital coherent optical receivers," *Opt. Commun.*, vol. 501, Dec. 2021, Art. no. 127380.
[5] F. N. Khan, K. Zhong, X. Zhou, W. H. Al-Arashi, C. Yu, C. Lu, and A. P. T. Lau, "Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks," *Opt. Exp.*, vol. 25, no. 15, pp. 17767–17776, Jul. 2017.
[6] P. Zhou, Y. Lu, D. Chen, and C. Li, "Relative entropy-based modulation format identification for coherent optical communication system," *Opt. Eng.*, vol. 60, no. 8, Aug. 2021, Art. no. 086101.
[7] L. Jiang, L. Yan, A. Yi, Y. Pan, M. Hao, W. Pan, and B. Luo, "An effective modulation format identification based on intensity profile features for digital coherent receivers," *J. Lightw. Technol.*, vol. 37, no. 19, pp. 5067–5075, Oct. 1, 2019.
[8] W. Zhang, J. Ye, Z. Yue, Y. Wang, X. Zhang, X. Zhang, and L. Xi, "Identifying modulation formats using integrated clustering algorithm," 2021, *arXiv:2104.14475*.
[9] A. Yi, L. Yan, H. Liu, L. Jiang, Y. Pan, B. Luo, and W. Pan, "Modulation format identification and OSNR monitoring using density distributions in Stokes axes for digital coherent receivers," *Opt. Exp.*, vol. 27, no. 4, pp. 4471–4479, 2019.
[10] J.-W. Kim and C.-H. Lee, "Modulation format identification of square and non-square M-QAM signals based on amplitude variance and OSNR," *Opt. Commun.*, vol. 474, Nov. 2020, Art. no. 126084.
[11] M. A. Jalil, J. Ayad, and H. J. Abdulkareem, "Modulation scheme identification based on artificial neural network algorithms for optical communication system," *J. ICT Res. Appl.*, vol. 14, no. 1, p. 69, Jul. 2020.
[12] V.-S. Doan, T. Huynh-The, C.-H. Hua, Q.-V. Pham, and D.-S. Kim, "Learning constellation map with deep CNN for accurate modulation recognition," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2020, pp. 1–6.
[13] F. N. Khan, Y. Zhou, A. P. T. Lau, and C. Lu, "Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks," *Opt. Exp.*, vol. 20, no. 11, pp. 12422–12431, May 2012.
[14] J. Du, T. Yang, X. Chen, J. Chai, Y. Zhao, and S. Shi, "A CNN-based cost-effective modulation format identification scheme by low-bandwidth direct detecting and low rate sampling for elastic optical networks," *Opt. Commun.*, vol. 471, Sep. 2020, Art. no. 126007.
[15] Y. Zhao, C. Shi, T. Yang, J. Du, Y. Zang, D. Wang, X. Chen, L. Wang, and Z. Zhang, "Low-complexity and joint modulation format identification and OSNR estimation using random forest for flexible coherent receivers," *Opt. Commun.*, vol. 457, Feb. 2020, Art. no. 124698.
[16] J. Hong, L. Chen, J. Zhu, W. Zhou, B. Li, Y. Fu, and L. Wang, "Modulation format identification and transmission quality monitoring for link establishment in optical network using machine learning techniques," in *Proc. Asia Commun. Photon. Conf./Int. Conf. Inf. Photon. Opt. Commun. (ACP/IPOC)*, 2020, pp. M4A–191.
[17] J. W. Kim and C. H. Lee, "Modulation format identification for square M-QAM signals by using a neural network," in *Proc. Asia Commun. Photon. Conf.*, 2017, pp. M1H–3.
[18] J. Zhang, Y. Li, S. Hu, W. Zhang, Z. Wan, Z. Yu, and K. Qiu, "Joint modulation format identification and OSNR monitoring using cascaded neural network with transfer learning," *IEEE Photon. J.*, vol. 13, no. 1, pp. 1–10, Feb. 2021.
[19] A. K. Ali and E. Erçelebi, "Algorithm for automatic recognition of PSK and QAM with unique classifier based on features and threshold levels," *ISA Trans.*, vol. 102, pp. 173–192, Jul. 2020.
[20] Y. Zhao, Z. Yu, Z. Wan, S. Hu, L. Shu, J. Zhang, and K. Xu, "Low complexity OSNR monitoring and modulation format identification based on binarized neural networks," *J. Lightw. Technol.*, vol. 38, no. 6, pp. 1314–1322, Mar. 15, 2020.
[21] C. Zhu, S. Han, H. Mao, and W. J. Dally, "Trained ternary quantization," 2016, *arXiv:1612.01064*.
[22] S. Yin, P. Ouyang, J. Yang, T. Lu, X. Li, L. Liu, and S. Wei, "An energy-efficient reconfigurable processor for binary-and ternary-weight neural networks with flexible data bit width," *IEEE J. Solid-State Circuits*, vol. 54, no. 4, pp. 1120–1136, Apr. 2018.
[23] A. K. Ali and E. Erçelebi, "Modulation format identification using supervised learning and high-dimensional features," *Arabian J. Sci. Eng.*, vol. 47, pp. 1–26, May 2022.
[24] Z. Zhao, A. Yang, and P. Guo, "A modulation format identification method based on information entropy analysis of received optical communication signal," *IEEE Access*, vol. 7, pp. 41492–41497, 2019.
[25] Y. Zhang, P. Zhou, C. Dong, Y. Lu, and L. Chuanqi, "Intelligent equally weighted multi-task learning for joint OSNR monitoring and modulation format identification," *Opt. Fiber Technol.*, vol. 71, Jul. 2022, Art. no. 102931.
[26] H. Lv, X. Zhou, J. Huo, and J. Yuan, "Joint OSNR monitoring and modulation format identification on signal amplitude histograms using convolutional neural network," *Opt. Fiber Technol.*, vol. 61, Jan. 2021, Art. no. 102455.

PENG ZHOU received the M.S. degree from Guangxi Normal University, China, in 2019, where he is currently pursuing the Ph.D. degree. He is also a member of the Laboratory of Optoelectronics and Optical Communications, China. His research interests include coherent optical communication systems, pattern recognition, and modulation format recognition.

CHUANQI LI received the B.E. degree from the Institute of Plasma, Chinese Academy of Sciences, China, in 1991, and the Ph.D. degree from Southeast University, China, in 2004. He is currently a Professor with Nanning Normal University, China. His current research interests include optical fiber communications and micro nano materials.

DONG CHEN received the M.S. degree from Harbin Engineering University, China, in 2009. He is currently pursuing the Ph.D. degree with Guangxi Normal University. His research interests include wireless network relay transmission technology and signal processing for optical transmission networks.

YU ZHANG received the B.S. degree from the Jiangsu Institute of Technology, China, in 2019. He is currently pursuing the M.S. degree with Guangxi Normal University. He is also a member of the Laboratory of Optoelectronics and Optical Communications, China. His research interests include coherent optical communication systems and deep learning.

YE LU received the M.S. degree from Guangxi Normal University, China, in 2015. He is currently a Lecturer with the College of Electronic Engineering, Guangxi Normal University. His research interests include codes for the OCDMA systems, coherent optical communication systems, and signal processing for optical transmission networks.