56-Gbit/s PAM-4 Optical Signal Transmission Over 100-km SMF Enabled by TCNN Regression Model

Jing Zhang, Lianshan Yan, *Senior Member, IEEE*, Lin Jiang, Anlin Yi, Yan Pan, Wei Pan, *Member, IEEE*, and Bin Luo, *Member, IEEE*

***Abstract*—Due to the interaction between chromatic dispersion (CD) and direct detection, the CD induced power-fading effect has been considered as the key point to limit transmission rate and distance in intensity modulation direct-detection (IM/DD) systems. Besides, the performance of the IM/DD systems will be further affected by the bandwidth limitation and nonlinearity effect. In this paper, we propose a low complexity nonlinear equalizer (NLE) based on temporal convolutional neural network (TCNN) regression model which combines dilated convolutions with residual connections. Then, the performance of our proposed equalizer is experimentally demonstrated in 56-Gbit/s 4-level pulse amplitude modulation (PAM4) intensity modulation direct-detection system. The results show that the receiver sensitivity with the help of our TCNN equalizer can be improved about 1.5 dB and 3.5 dB compared to 2D-CNN and 1D-CNN equalizer under 70-km single mode fiber (SMF) transmission. Furthermore, the proposed equalizer can extend maximum transmission distance to 100 km at bit error rate threshold of 3.8 × 10^{-3}.**

***Index Terms*—Convolutional neural network, direct detection, intensity modulation, short-reach transmission link.**

I. Introduction

DRIVEN by the growing demands for bandwidth-consuming applications (e.g., cloud computing, Internet of Things (IOT) and artificial intelligence-intensive services and so on), higher-speed access network would be required for meeting the increasing data traffic. Compared with other mainstream technology solutions applied in access network, intensity-modulation and direct-detection (IM/DD) with low cost and power consumption has obtained a lot of attentions [1]–[3]. Usually, the single-end photodetector which has been applied to detect intensity information is considered as the key receiver component of IM/DD systems. Here, in order to improve system capacity, various advanced intensity modulation formats are introduced, such as discrete multi-tone (DMT) [4], carrier-less amplitude phase (CAP) [5], pulse-amplitude modulation (PAM) [6] and Nyquist sub-carrier modulation (N-SCM) [7]. Among these, PAM-4 with the low power consumption and implementation complexity has gained more popularity compared to other modulation formats in IM/DD systems. Certainly, IM/DD systems may be more sensitive to various linear and nonlinear impairments with the applications of advanced intensity modulation formats.

In IM/DD systems, due to the interaction between chromatic dispersion (CD) and direct detection, the CD induced power-fading effect may be a major obstacle to limit the achievable capacity-distance product. Here, various effective technologies have been introduced to alleviate CD-induced spectrum null, including dispersion compensation fiber, single-sideband or vestigial-sideband (SSB/VSB) modulation and so on. However, these techniques may increase the implementation complexity and cost through additional device or technology assistance. Therefore, advanced digital signal processing (DSP) algorithms in the receiver may be a better alternative which would not change the structure of IM/DD systems. In recent years, various DSP algorithms have been reported in IM/DD systems to alleviate or mitigate CD effect, such as feed-forward equalization (FFE), decision feedback equalization (DFE), Volterra equalizer (VE), maximum likelihood sequence estimation (MLSE), Tomlinson-Harashima pre-coding and neural network. Among these algorithms, neural network with the inherent advantage of approximating any nonlinear function is considered to have great potential to mitigate CD effect and device nonlinear effect in IM/DD systems [8]–[11]. The development of integrated photonics technology [12] offers great potential for the application of neural networks. It is well known that the default choice for sequence modeling task was recurrent neural networks (RNNs) because of their great ability to capture temporal dependencies in sequential data. However, due to the recurrent structure in RNN model, the computation of one time point must wait for the result of the former time point, so RNNs cannot perform massively parallel processing like convolutional neural networks (CNNs) resulting in huge time-consumption.

In this paper, we have proposed a low complexity nonlinear equalizer (NLE) based on temporal convolutional neural network (TCNN) with an excellent ability of impairment equalization which is an extension of our previous work [13]. The proposed TCNN equalizer is a regression model which is developed to take complementary advantage of parallel computing capability of CNNs and temporal sensitivity of RNNs. The

Manuscript received May 6, 2021; revised June 15, 2021; accepted June 21, 2021. Date of publication June 24, 2021; date of current version August 10, 2021. This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFB1803500, and in part by National Natural Science Foundation of China (NSFC) under Grants 61860206006 and 62005228. *(Corresponding author: Lin Jiang.)*

The authors are with the Center for Information Photonics & Communications, School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China (e-mail: jingzhang@my.swjtu.edu.cn; lsyan@swjtu.edu.cn; linjiang@swjtu.edu.cn; anlinyi@swjtu.edu.cn; py_swjtu@my.swjtu.edu.cn; wpan@home.swjtu.edu.cn; bluo@home.swjtu.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2021.3092003

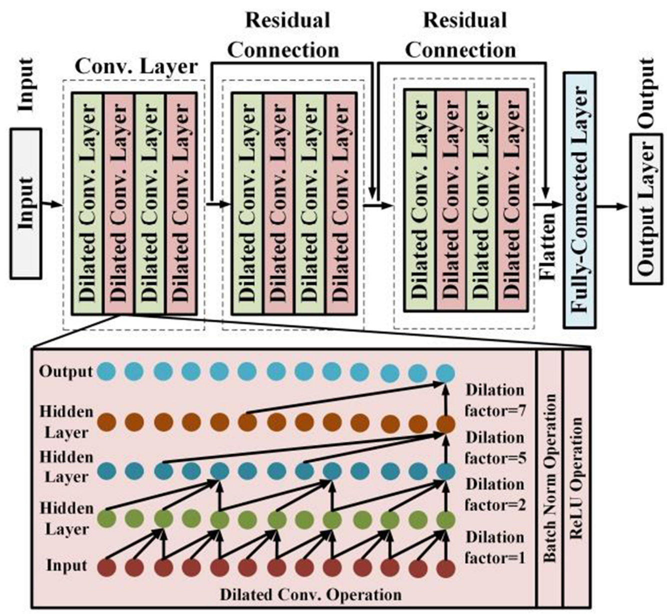

Fig. 1. Structure of the proposed TCNN equalizer.

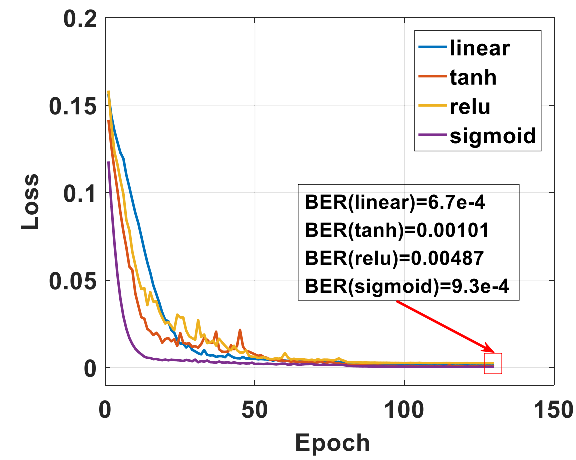

Fig. 2. The loss value of different activation functions.

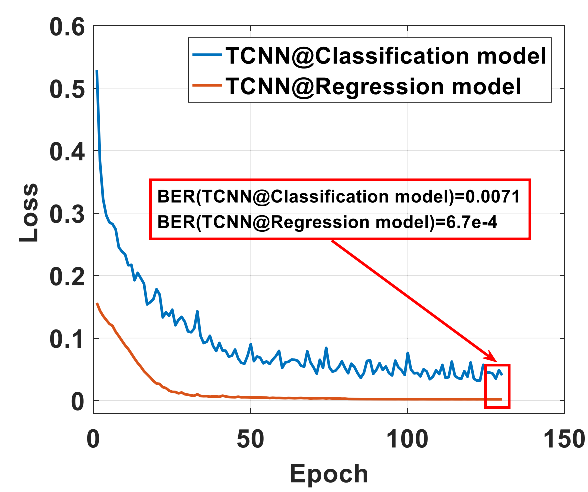

Fig. 3. The loss value of TCNN equalizer with regression model and classification model under different epochs.

performance of the proposed TCNN equalizer has been experimentally demonstrated in a 56-Gbit/s PAM-4 IM/DD dispersion-uncompensated system. The results show that the receiver sensitivity with the help of our proposed TCNN equalizer can be improved about 1.5 dB and 3.5 dB compared to two-dimensional (2D) CNN and one-dimensional (1D) CNN equalizer under 70-km single mode fiber (SMF) transmission. Furthermore, the proposed TCNN equalizer can extend maximum transmission distance to 100 km at bit error rate threshold of 3.8×10^{-3}.

II. Operation Principle

Fig. 1 shows the structure of the proposed TCNN equalizer which contains an input layer with 2x51 array for TCNN input, 12 convolutional layers with 64 channels, a fully connected layer with 256 neuron nodes and an output layer with one neuron node. The "channel" in the proposed TCNN equalizer represents the dimensionality of the output space. In our TCNN equalizer, rectified linear units (ReLU) function is considered as the activation function of convolutional layer and fully connected layer, while linear activation function is applied in the output layer. Unlike most neural network-based equalization algorithms which treat the problem of impairments equalization as signal classification problem where one-hot coding is usually used to represent the amplitude levels, the signal impairments equalization is regarded as a regression problem in our TCNN equalization algorithm and the model output is accurately predicted by statistical analysis. For regression model, it is reasonable to choose a linear and sigmoid output activation function. Subsequently, we have added some comparative test of different activation functions as shown in Fig. 2. We find that the linear output activation function can present the best performance in our TCNN model.

The fundamental difference between regression and classification is whether the output space is a metric space. The regression model defines a metric formula MSE (Mean Square Error) to measure the error between the output and the label. When the input symbol 1 is predicted to be symbol 2, the error is 1. When the input symbol 1 is predicted to be symbol 3, the error is 4. Therefore, the regression model can distinguish the above two different situations. However, the output space of classification model is not a metric space, which is so-called "qualitative". That is to say, there is only the distinction between "correct" and "error" in the classification model. As for whether to classify symbol 1 into symbol 2, or symbol 3, there is no difference.

Compared to classification model with 4 outputs, the regression model with 1 output shows better performance as depicted in Fig. 3. In addition, the regression model in this paper can save more computational complexity since 3 output neurons are waived. Besides, compared with traditional CNN, TCNN introduced dilated convolutions and residual connections. It is worth noting that the premise of not losing useful information depends on the design of the dilated convolution. Dilated convolutions are a type of convolution that "inflate" the kernel by inserting holes between the kernel elements. In addition, the employment of residual connections can ensure the stability of deep TCNN model.

A. Input Sequence

The received optical signal is converted from optical domain to electric domain by a single-end photodetector (PD), and then the obtained electric signal are sampled into digital signal by analog digital converter (ADC). The up-sampled digital signal with 2 samples per symbol is fed into the proposed TCNN equalizer. The input serial sequences are firstly transformed into a series of parallel sequences with 2×51 data array, and the difference between the two rows is only one sample. These parallel sequences can be used as the 2D input data of the first convolutional layer. To ensure the fairness of the comparison, the total amount of training data is the same for each neural network-based equalizer.

B. Dilated Convolutions and Residual Connections

Compared with traditional convolutions operation, the kernel of dilated convolution introduces "holes" to the traditional convolutions, which enable an exponentially large receptive field without losing resolution or coverage [14]. Moreover, it can solve the problem of down-sampling which result in loss of information. Usually, for 1-D input sequence $\mathbf{X}\in\mathbb{R}^{n}$ and a convolution kernel $\mathbf{f}$: $\{0, 1, 2, 3, \ldots, k\text{-}1\}\rightarrow\mathbb{R}$, the dilated convolution operation F can be defined as:

$$\mathrm{F}(s) = (\mathrm{X} *_{d} \mathrm{f})(s) = \sum_{i=0}^{k-1} \mathrm{f}(i) \cdot \mathrm{X}_{s-d \cdot i} \tag{1}$$

where d is the dilation factor, which refers to the number of intervals between two adjacent points in the convolution operation. s is the element of the input sequence. k is the convolution kernel size. When the dilation factor $d = 1$, the dilated convolution reduces to a traditional convolution. There are two ways to increase the receptive field of the TCNN. The one is to increase the dilation factor d, and the other one is to choose a larger convolution kernel $\mathbf{f}$. In our proposed TCNN equalizer, we gradually increase the dilation factor d for improving the receptive field. The convolution kernel size is fixed as (1, 3) with less training parameters of the model.

The emergence of residual blocks solves the vanishing gradient problem on deep convolutional neural network. There are two residual blocks which are applied in our TCNN equalizer. Each residual block has a main connection and short connection. The main connection contains four convolutional layers to extract the characteristic parameters. The residual block performs a series of transformations F on the input $\mathbf{x}$, whose outputs are added to the input $\mathbf{x}$ of the block:

$$\mathrm{Output} = \mathrm{Activation}(\mathbf{x} + F(\mathbf{x})) \tag{2}$$

Residual connection solves gradient vanishing/explosion and network degradation problems in deep neural network.

C. Model Training

In our TCNN equalizer, Adam optimization algorithm [15] is applied for minimizing the mean squared error between the predicted value $\mathbf{Y_p}$ and the true value $\mathbf{Y_t}$:

$$L(\mathrm{W}, \mathrm{b}) = \|\mathrm{Y_t} - \mathrm{Y_p}\|_2^2 \tag{3}$$

where $\mathbf{W}$ is weight and $\mathbf{b}$ is bias. For further improving Adam's performance, we combine Adam optimization and learning rate decay to update parameters of network. The initial learning rate is set 0.0001$\sim$0.001, and the learning rate is reduced to 0.1 times of the original when the loss no longer drops. To avoid over-fitting and improve training performance, we add the batch normalization after each layer.

III. Experimental Setup

The experimental setup of 56-Gbit/s PAM-4 transmission system is depicted in Fig. 4. It is well known that PRBS is not suitable for the research of equalization based on neural network, since neural network with powerful recognizing and modeling capabilities has the potential to characterize the generation rule of PRBS. Once PRBS is used as training set and test set, the performance of neural network-based equalizer may be overestimated [16], [17]. In our experiment, the data sets used to train and test are all random sequences generated by mixing multiple random sequences.

At transmitter side, the light from laser at $\sim$1549.32 nm with $\sim$100-kHz linewidth is modulated by a 40-GHz Mach-Zehnder modulator (MZM) driven by the DAC operation 64-GSa/s and 25-GHz analog bandwidth to generate 56-Gbit/s PAM-4 optical signal. Subsequently, the signal optical power after MZM modulator needs to be adjusted by an erbium doped fiber amplifier (EDFA) and a variable optical attenuator (VOA). And then, the optical signal is launched into a dispersion-uncompensated transmission link which consists of multiple fiber spans.

At receiver side, a VOA followed by a second EDFA is employed to adjust the received optical power (ROP) to a certain level. The received optical signal is converted to an electrical signal by a 17-GHz PD. Finally, the received electrical signal is digitized by an 80-GSa/s and 33-GHz analog bandwidth digital storage oscilloscope (DSOZ634A). Noted that 100,000 samples are considered as training set, another 100,000 samples are considered as testing set. In the offline DSP, the digital waveform is first resampled to 4-sps, and then passed through a matched filter to filter out the out-of-band noise. Then, in order to ensure the reasonableness of the comparison, the one dimensional (1D)-CNN, two dimensional (2D)-CNN, fully connected neural network (FCNN) and Volterra are introduced to measure the equalization performance. And all models are designed according to the optimal parameters.

IV. Experimental Results and Discussion

In this section, we present the experimental results of four equalizers under different conditions (e.g., different received powers, different launch powers, and different transmission distances). Firstly, the BER results under different received powers

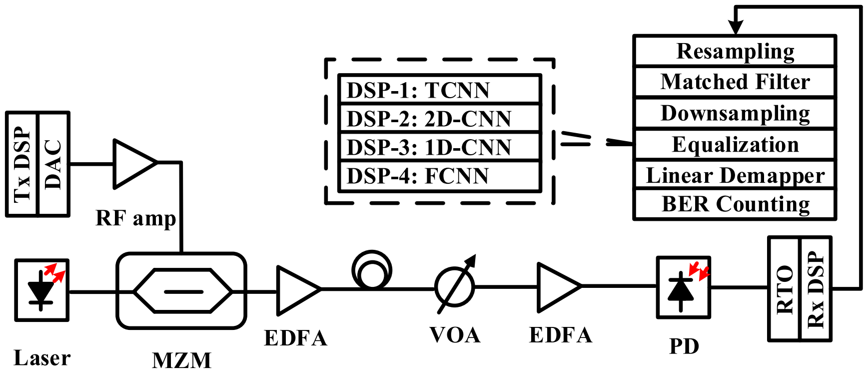

Fig. 4. Experimental setup of the 56-Gbit/s PAM4 transmission system.

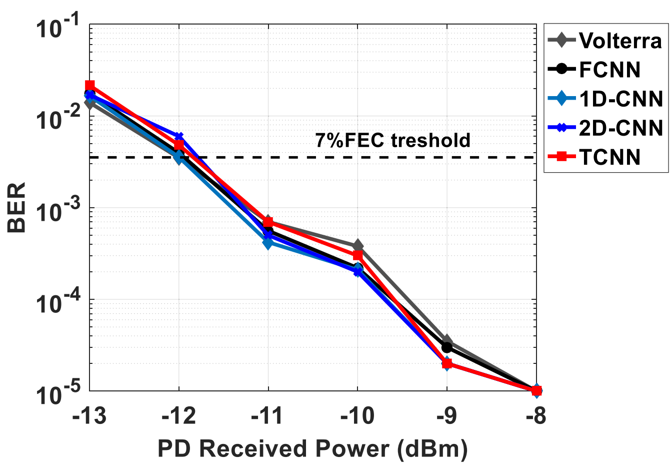

Fig. 5. BER performance with different PD received powers at back-to-back transmission link.

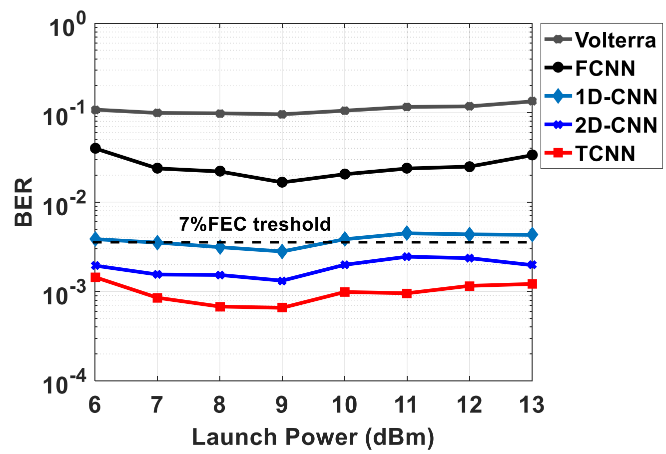

Fig. 6. BER performance with different launch powers at 70-km SSMF transmission link.

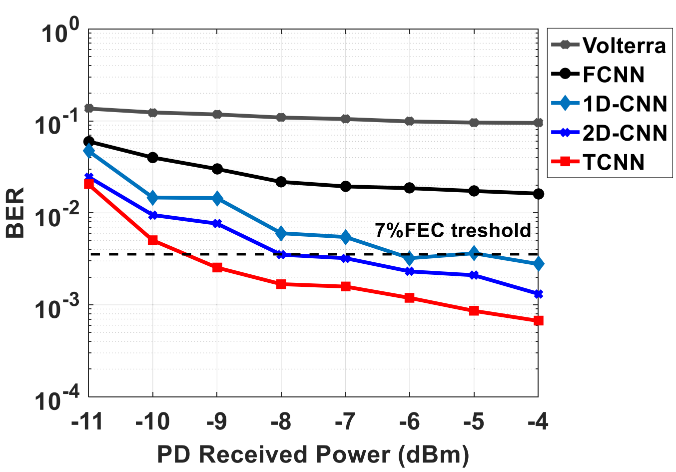

Fig. 7. BER performance with different PD received powers at 70-km SSMF transmission link.

ranging from -13 dBm to -8 dBm for back-to-back case are depicted in Fig. 5. Since optical signal may be interfered by the single-end PD noise including thermal noise and shot noise, the BER performance would decrease as the injected received power decreases for all NN-based equalizers. It is well known that all NN-based equalizers or other impairments equalization algorithms cannot eliminate the thermal noise and shot noise of PD, so that the BER performance of our proposed TCNN equalizer is the same as other equalizers.

In short-reach optical fiber communication systems, fiber nonlinearity is indeed not the primary consideration. In optical access networks, there will be a situation where one optical line terminal (OLT) connects multiple optical network units (ONUs), so that the launch power is still quite large. In this case, the fiber nonlinearity can't be ignored. Fig. 6 shows the BER performance under different launch powers (e.g., ranging from 6 dBm to 13 dBm) at 70-km SSMF transmission link with -4-dBm received power. The proposed TCNN, 1D-CNN, 2D-CNN, FCNN and Volterra, all have the same optimum launch power of 9 dBm, while the proposed TCNN equalizer with powerful nonlinearity equalization ability obtains the best BER performance compared with other equalizers. When the launch power is larger than 9 dBm, the enhancement of nonlinear effects has exceeded the equalization capability of these equalizers, which would result in severe performance degradation. Therefore, we keep the launch power to be 9 dBm for each case in the following experiments.

Fig. 7 shows the BER curves as a function of different received powers at 56-Gbit/s PAM-470-km SSMF transmission links.

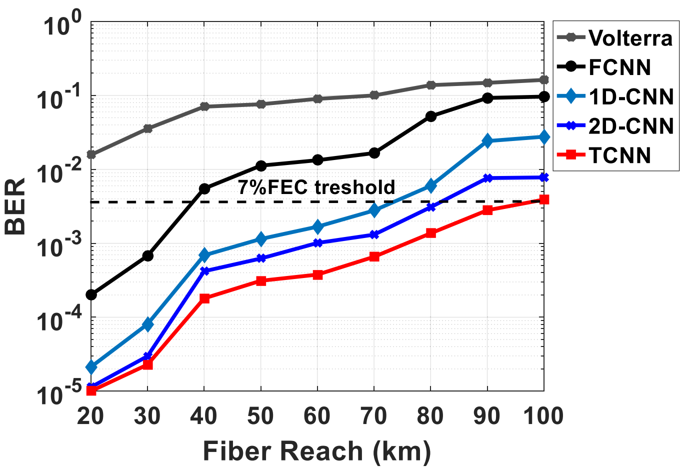

Fig. 8. BER performance with different transmission distances at 9-dBm launch power.

As expected, the BER performance of all equalizers decrease as the injected received power decreases. Here, the proposed TCNN equalizer, 2D-CNN equalizer, and 1D-CNN equalizer can achieve the receiver sensitivity of approximately −9.5 dBm, −8 dBm and −6 dBm at 7% FEC threshold (BER = 3.8 × 10-3), respectively. We find that the FCNN never reaches the 7% FEC threshold even with the highest received optical power over 70-km transmission link. The receiver sensitivity with the help of TCNN equalizer can be improved about 1.5 dB and 3.5 dB compared to 2D-CNN and 1D-CNN equalizer.

At last, for further investigating the performance of the proposed equalizer, we carried out a series of transmission experiments under different transmission distances. Fig. 8 shows that the proposed TCNN equalizer can perform the best performance in the presence of a large amount of dispersion than the other four equalizers. Given 7% FEC threshold (BER = 3.8×10^{-3}), the transmission system assisted by 2D-CNN equalizer, 1D-CNN equalizer and FCNN equalizer just realize the maximum transmission distance of 80 km, 72 km and 38 km, respectively. Here, the proposed TCNN equalizer can extend maximum transmission distance to 100 km.

The complexity of the model is measured by the number of multiplications required to run once [18]. The mathematical expression of 3-order Volterra with memory length of L_i can be expressed as:

$$h(n) = \sum_{i=0}^{L_1-1} w_i x(n-i) + \sum_{i=0}^{L_2-1}\sum_{j=0}^{i} w_{i,j} x(n-i)x(n-j) + \sum_{i=0}^{L_3-1}\sum_{j=0}^{i}\sum_{k=0}^{j} w_{i,j,k} x(n-i)x(n-j)x(n-k) \quad (4)$$

Where *x(n)* is nth sampled data from received signals and *h(n)* is output data through equalization. **W** is Volterra kernel. The tap numbers of three kernels are *L1*, *L2*(*L2*+1)/2, *L3*(*L3*+1)(*L3*+2)/6, respectively. Therefore, the computation complexity of three-order Volterra is *L1*+2 × *L2*(*L2*+1)/2+3 × *L3*(*L3*+1)(*L3*+2)/6. Therefore, the number of multiplications of Volterra (91, 31, 7) is 1335.

The computation complexity for fully connected neural networks is defined as number of multiplication operation as calculated [9]:

$$Complexity\_full\_connected = N1 \times N2 + N3 + \cdots \quad (5)$$

where *N1*, *N2*, and *N3* are the neuron number of input, hidden and output layer, respectively. In addition, for a convolutional neural network, the computational complexity is the sum of the complexity of all fully connected layers as defined in (5) and the complexity of all convolutional layers which can be defined as [19]:

$$Complexity\_convolutional = N_{cin} \times K_w \times K_h \times N_{cout} \times H_o \times W_o \quad (6)$$

where N_{Cin} represents the number of channels of the input data, K_w and K_h represent the width and height of the convolution kernel, N_{Cout} represents the number of channels of the output data, and the H_O and W_O represent the height and width of the output data. The TCNN model in this paper consists of an input layer with 2x51 array, a fully connected layer with 256 neurons, 12 convolutional layers with 64 channels whose kernel size is 1 × 3, and an output layer with one neuron. The number of multiplications of TCNN can be calculated as:

$$2 \times 1 \times 3 \times 64 \times 1 \times 51 + (64 \times 1 \times 3 \times 64 \times 1 \times 51) \times 11 + 64 \times 51 \times 256 + 256 \times 1 = 7748992. \quad (7)$$

TABLE I
THE SUMMARY OF THE COMPUTATION COMPLEXITY

Model	Number of Multiplications
Volterra	1335
FCNN	22,3744
1D-CNN	2,8802,2528
2D-CNN	1,0403,6992
TCNN	774,8992

Similarly, the complexity of other models can be calculated. The relevant quantitative analysis of complexity has been given as Table I. The computation complexity from high to low can be arranged as 1D-CNN > 2D-CNN> TCNN > FCNN> Volterra. Although the complexity of Volterra is the lowest, the equalization performance of TCNN is better than that of Volterra. Therefore, we believe that TCNN equalizer can be regarded as a good candidate after the tradeoff between complexity and performance.

V. CONCLUSION

In this paper, we propose a low complexity nonlinear equalizer (NLE) based on temporal convolutional neural network (TCNN) regression model. This model has excellent impairment equalization ability and has been experimentally demonstrated in C-band 56-Gbit/s PAM4 IM/DD optical system over 100-km dispersion-uncompensated link. The proposed TCNN equalizer can compensate most of linear and nonlinear distortions. In addition, we design a relatively random data set as training and testing set to prevent overestimating of NN-based equalizer in performance evaluation process. Compared with the previous equalizers such as 1D-CNN equalizer, 2D-CNN equalizer, FCNN equalizer and Volterra, the proposed TCNN equalizer shows the best equalization performance. Results show that the TCNN equalizer achieves a 1.5 dB and 3.5 dB better sensitivity than 2D-CNN and 1D-CNN respectively under 70-km SSMF transmission.

Acknowledgment

The authors wish to thank the anonymous reviewers for their valuable suggestions.

References

[1] W. Yan *et al.*, "80 km IM-DD transmission for 100 Gb/s per lane enabled by DMT and nonlinearity management," in *Proc. OFC*, San Francisco, CA, USA, 2014, pp. 1–3.

[2] X. Tang, S. Liu, Z. Sun, H. Cui, and Y. Qiao, "C-band 56-Gb/s PAM4 transmission over 80-km SSMF with electrical equalization at receiver," *Opt. Exp.*, vol. 27, pp. 25708–25717, 2019.

[3] H. Wang *et al.*, "Adaptive channel-matched detection for C-Band 64-Gbit/s optical OOK system over 100-km dispersion-uncompensated link," *IEEE/OSA J. Lightw. Technol.*, vol. 38, no. 18, pp. 5048–5055, Sep. 2020.

[4] L. sun, J. Du, C. Wang, Z. Li, K. Xu, and Z. He, "Frequency-resolved adaptive probabilistic shaping for DMT-modulated IM-DD optical interconnects," *Opt. Exp.*, vol. 27, pp. 12241–12254, 2019.

[5] L. Tao, Y. Wang, J. Xiao, and N. Chi, "Enhanced performance of 400 Gb/s DML-based CAP systems using optical filtering technique for short reach communication," *Opt. Exp.*, vol. 22, pp. 29331–29339, 2014.

[6] D. Sadot, G. Dorman, A. Gorshtein, E. Sonkin, and O. Vidal, "Single channel 112 Gbit/sec PAM4 at 56 Gbaud with digital signal processing for data centers applications," *Opt. Exp.*, vol. 23, pp. 991–997, 2015.

[7] Z. Li *et al.*, "Signal-signal beat interference cancellation in spectrally-efficient WDM direct-detection Nyquist-pulse-shaped 16-QAM subcarrier modulation," *Opt. Exp.*, vol. 23, pp. 23694–23709, 2015.

[8] A. Reza and J. Rhee, "Nonlinear equalizer based on neural networks for PAM-4 signal transmission using DML," *IEEE Photon. Technol. Lett.*, vol. 30, no. 15, pp. 1416–1419, Aug. 2018.

[9] Z. Xu, C. Sun, T. Ji, J. Manton, and W. Shieh, "Computational complexity comparison of feedforward/radial basis function/recurrent neural network-based equalizer for a 50-Gb/s PAM4 direct-detection optical link," *Opt. Exp.*, vol. 27, pp. 36953–36964, 2019.

[10] Z. Xu, C. Sun, T. Ji, J. Manton, and W. Shieh, "Cascade recurrent neural network-assisted nonlinear equalization for a 100 Gb/s PAM4 short-reach direct detection system," *Opt. Lett.*, vol. 45, pp. 4216–4219, 2020.

[11] X. Dai, X. Li, M. Luo, Q. You, and S. Yu, "LSTM networks enabled nonlinear equalization in 50-Gb/s PAM-4 transmission links," *Appl. Opt.*, vol. 58, no. 22, pp. 6079–6084, 2019.

[12] D. Liang and J. Bowers, "Recent progress in heterogeneous III-V-on-silicon photonic integration," *Light Adv. Manuf.*, vol. 2, 2021, doi:10.37188/lam.2021.005.

[13] J. Zhang *et al.*, "Convolutional neural network equalizer for Short-reach optical communication systems," in *Proc. Asia Commun. Photon. Conf. (ACP) Int. Conf. Inf. Photon. Opt. Commun.*, Beijing, China, 2020, Paper M4A.320.

[14] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2016, *arXiv:1511.07122*.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980, 2014, [Online] Available: http://arxiv.org/abs/1412.6980.

[16] L. Yi, T. Liao, L. Huang, and L. Xue, "Machine learning for 100 Gb/s/λ passive optical network," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1621–1630, Mar. 2019.

[17] T. A. Eriksson, H. Bülow, and A. Leven. "Applying neural networks in optical communication systems: Possible pitfalls," *IEEE Photon. Technol. Lett.*, vol. 29, no. 23, pp. 2091–2094, Dec. 2017.

[18] B. S. G. Pillai *et al.*, "End-to-end energy modeling and analysis of long-haul coherent transmission systems," *J. Lightw. Technol.*, vol. 32, no. 18, pp. 3093–3111, Sep. 2015.

[19] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 5353–5360, doi: 10.1109/CVPR.2015.7299173.