Improved CNN Equalizer for Coherent Optical Fiber Communications

1st Lu Han
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
clear@bupt.edu.cn

2nd Yongjun Wang*
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
*wangyj@bupt.edu.cn

3rd Chao Li
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
lythao@bupt.edu.cn

4th Xingyuan Huang
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
yuan0624@bupt.edu.cn

5th Qi Zhang
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
zhangqi@bupt.edu.cn

6th Heming Zhu
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
2021110517_zhm@bupt.edu.cn

7th Qinghua Tian
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
tianqh@bupt.edu.cn

8th Xiangjun Xin
Beijing University of Posts and Telecommunications (BUPT)
State Key Laboratory of Information Photonics
and Optical Communications(IPOC)
Beijing, China
xjxin@bupt.edu.cn

***Abstract*—We propose an improved CNN equalizer with large kernel designment. We achieve ∼1.58 dB Q-factor improvement for 120 Gb/s PDM 64-QAM transmission over 375 km transmission distance and ∼27% time complexity reduction.**

***Index Terms*—Large kernel, channel equalization, complex reduction**

I. Introduction

Driven by the Internet of Things service, the demand for high data capacity and high-speed data transmission has stimulated the research and development of ultra-high-speed, large-capacity optical transmission and highly flexible optical network technology [1]. To cope with the explosive growth of data service demand, optical fiber communication system adopt various multiplexing methods like polarization division multiplexing (PDM), wavelength-division multiplexing (WDM), space division multiplexing (SDM) and high-order quadrature amplitude modulation (QAM) [2]. However, nonlinearity caused by Kerr effects from fiber transmission link degrades the system quality [3]. Various nonlinear damage compensation methods, including DBP based methods [4], perturbation theory based methods [5], and machine learning based methods [6,7], have been proposed successively. In our previous work [8], a perturbation theory aided convolutional neural network (CNN) equalizer has been proposed for PDM 64-QAM coherent detection, which has demonstrated the great potential of the joint processing of physical methods and machine learning. Recently, the scholars have demonstrated the big kernel can play an important role in CNN [9,10]. In [9], the work from Tsinghua University and Kuangshi Technology show that the performance of proposed RepLKNet of super kernel outperforms Swin and other transformers models. Scholars also found that the large kernel design is more like human because it has high shape bias, which means it extract information from the whole respective field, and get more edge information, not just focus on the central detail texture [10,11].

In this paper, we propose an improved CNN equalizer with

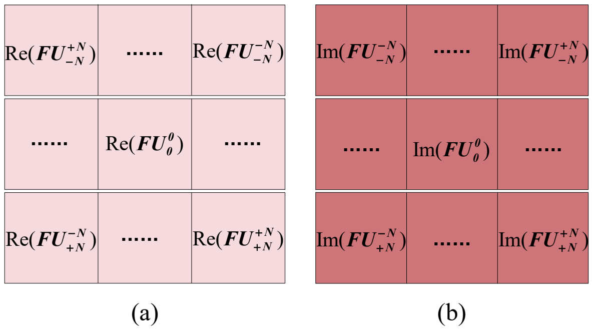

Fig. 1. Feature map Construction.

large kernel for nonlinear compensation in coherent optical fiber communication systems. We analyze the affect of feature map size, and select the optimum processing parameters. Then based on the optimum parameters, we compared the performance of large kernel aided classifier (LK-CNNC), the regressor (LK-CNNR), the small kernel CNN classifier (SK-CNNC), and the small kernel CNN regressor (SK-CNNR) at the same time complexity. The rest of our paper is organized as follows. In Section II, we discuss the basic principle of perturbation theory, and the improved LK-CNN. In Section III, the configuration of the coherent optical fiber communication system is presented. In section IV, we discuss the experimental results and discuss the system performance measured by Q-factor. Finally, we conclude in Section V.

II. Principles

In this section, we introduce the basic principle of perturbation theory, and we propose a large kernel aided CNN structure for nonlinearity compensation.

A. Feature Map Construction

In PDM optical fiber transmission system, equation characterizing the propagation of optical pulse can be expressed as the Manakov equation, and in X polarization, the Manakov equation can be expressed as following[5]:

$$\frac{\partial \psi_x}{\partial z} + \frac{\alpha}{2}\psi_x + i\frac{\beta_2}{2}\frac{\partial^2 \psi_x}{\partial t^2} = i\frac{8}{9}\gamma\left(|\psi_x|^2 + |\psi_y|^2\right)\psi_x, \quad (1)$$

where $\psi_x = \sum_s T_x[s] f(0, t - sT)$ refers to the optical carrier signals, $T_x[s]$ denotes the amplitude of the s_{th} code element in X polarization, T refers to the duration of the code element, and f denotes the waveform of a carrier pulse. α, β_2 and γ refer to the linear loss in the fiber optic links, group velocity dispersion coefficient and nonlinear Kerr coefficient, respectively.

Based on the large chromatic dispersion assumption, the perturbation Δu between transmitter side symbol $T_{x/y}[s]$ and receiver side symbol $R_{x/y}[s]$ can be expressed as the vector dot product of FU and the perturbation coefficient C_{mn}, which can be simplified as following equation:

$$\Delta u = R_{x/y}[s] - T_{x/y}[s] = \sum_{m,n} FU_m^n C_{m,n}, \quad (2)$$

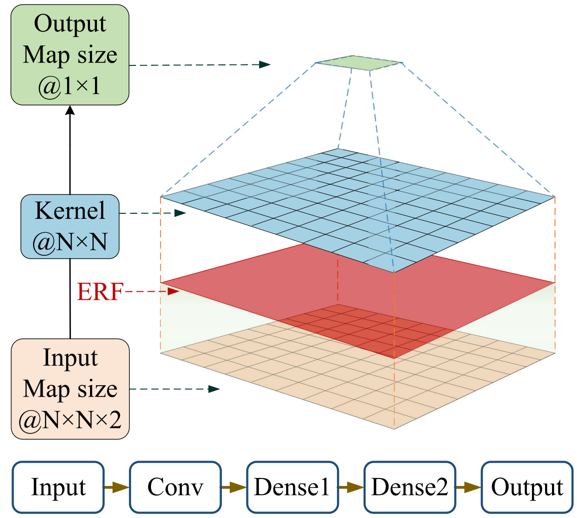

Fig. 2. Diagram of the ERF of large convolutional kernel.

where the perturbation coefficient $C_{m,n}$ can be calculated by the system link parameters. FU denotes the feature unit, and it can be expressed as the following equation [8]:

$$\begin{aligned} FU_m^n = & T_x[s+m]\left(T_x[s+n]T_x^*[s+m+n]\right) \\ & + T_x[s+m]\left(T_y[s+n]T_y^*[s+m+n]\right). \end{aligned} \quad (3)$$

We have constructed this dual-channel feature map shown in Fig. 1 and demonstrated its superiority in [8]. In subsequent experiments, we use this dual-channel feature map as the input image for further research.

B. Large-kernel CNN Design

The size of convolution kernel can be called as effective receptive field (ERF) which refers to the range of input figure processed simultaneously. As shown in Fig. 2, we used a convolution kernel structure that is equivalent to the size of the feature map to build the system architecture, it can be seen the ERF contains the entire feature map. The input feature map size is $N \times N \times 2$, and the kernel size is $N \times N$. After the convolution layer, the output size becomes 1×1, then the array has been flattened into a one-dimensional plane. The intermediate information transmission can be completed by a fully connected neural network. When the loss function is cross-entropy, the LK-CNN is acted as the classifier (LK-CNNC), and when the loss function is mean square error (MSE), the LK-CNN is acted as the regressor (LK-CNNR).

For the time complexity of equalizers, we generally use the required number of floating-point operations (FLOPs) when equalize one symbol. In the subsequent experimental processing, we has verified the system performance at the same time complexity. And it is shown that the large kernel structure can achieve better results than the small kernel cascade system structure.

III. Experimental Setup

The experimental set up for the 120 Gb/s PDM 64-QAM coherent optical communication system is shown in Fig. 3. At

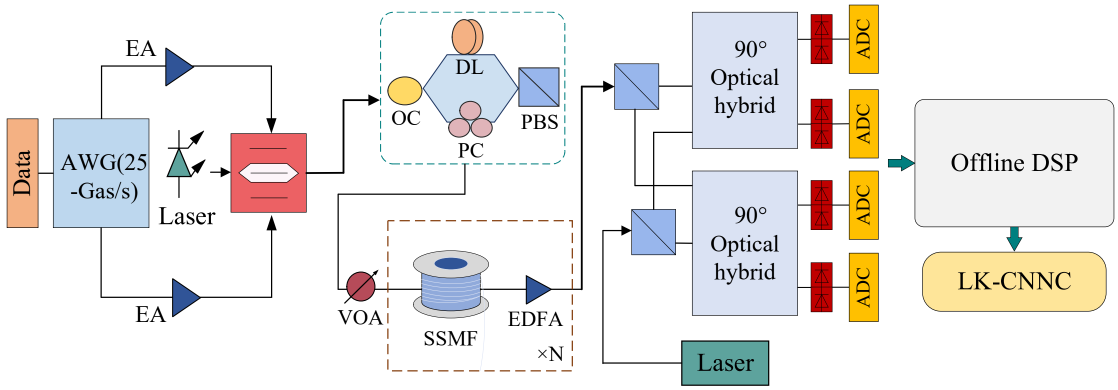

Fig. 3. Experimental Setup

the transmitter side, the 64-QAM mapping data are converted to electrical signals by arbitrary waveform generator (AWG) of 25 GSa/s sampling rate, and then amplified by two electric amplifier (EA). The external cavity laser (ECL) and the two electric signals are transmitted into the I/Q modulator to complete modulation. The PDM module is composed by polarization maintaining optical fiber coupler (PM-OC), optical delay line (DL), polarization controller (PC) and polarization beam combiner (PBC). The transmission link is 375 km G.652D single-mode fiber (SMF) of 5 span, and at each span, an EDFA with 20 dB gain is used to amplify the optical signal. At the receiver side, the optical signal and a local oscillator are sent to polarization beam splitter, then they are detected by 90-degree optical hybrid and balanced photonic detectors (BPDs). A real-time oscilloscope with 100 GSa/s sampling rate is used as an analog-to-digital (A/D) converter to acquire data. The offline DSP consists of low pass filter, I/Q imbalance compensation, CD compensation, clock recovery, polarization demultiplexing, PMD compensation, frequency offset estimation, carrier phase recovery and proposed LK-CNNC nonlinear equalization.

The LK-CNN is built, trained and evaluated in Pytorch 3.8.1 platform. The personal computer platform owns an AMD Ryzen7 CPU @ 2.90 GHz, and the Random Access Memory (RAM) is 16 GB. In our model, the Adam optimizer is employed to optimize the network. The datasets for each launched optical power (LOP) contain approximately 2^{20} symbols, and we divided them into 70% for training and 30% for testing, and the initial learning rate is set to 0.003.

IV. Results and Analysis

In this section, we analyze and select the appropriate size of feature map for higher performance. The Q-factor performance and time complexity of LK-CNNC are also compared with LK-CNNR, SK-CNNC and SK-CNNR. In our LK-CNNC structure, a convolution kernel that is equivalent to the size of the feature map is used to build the system architecture. As shown in Fig. 4, we analyze various feature map size ranged from 5 to 13 at LOP is 1 dBm, nonlinearity equalization capability is assessed by the Q-factor (the orange line), and the time complexity (the green blocks). The Q-factor is calculated from the bit error rate (BER), and can be calculated as:

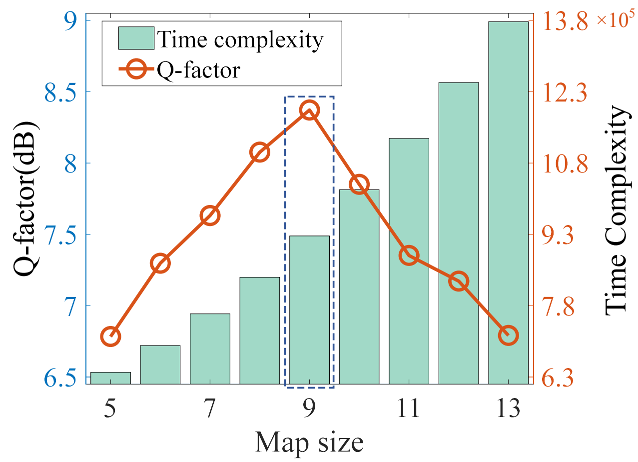

Fig. 4. Q-factor performance and time complexity of LK-CCNNC with different feature-map sizes at LOP is 1 dBm.

$$Q = 20\log_{10}\left(\sqrt{2}erfcinv\left(2 * BER\right)\right), \tag{4}$$

With the increase of the size of feature map, the Q-factor shows a trend of first rising and then falling, and the Q-factor get the maximum value when feature map size is 9. When the feature map size is small, the amplitudes of all perturbation coefficients are large, that is, all FUs are important, and contains more valid information. However, when a certain threshold is reached, with the increase of valid information, the additional noise is also increased, so the Q-factor is reduced. The time complexity has been increasing due to the increasing of the feature map size. Considering the Q-factor performance and time complexity, we choose 9 as the size of the optimal feature map, and the convolution kernel size of LK-CNN is also set 9 at subsequent experiments.

Based on the optimal parameters, we selected LK-CNNR, SK-CNNC, and SK-CNNR to compare the equalization performance and the time complexity. As shown in Fig. 5, it plots

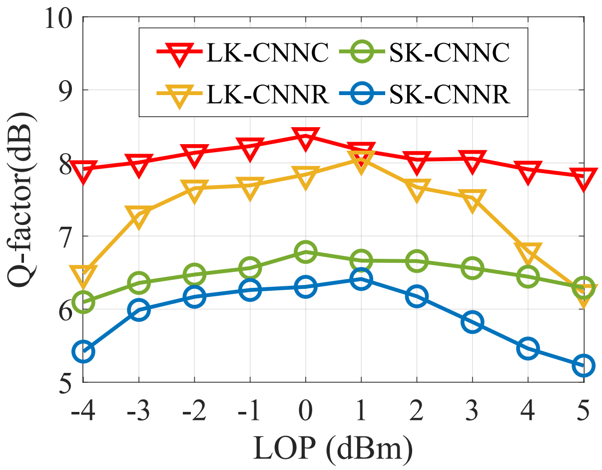

Fig. 5. Nonlinear equalization performance of different neural networks with the same time complexity.

the Q-factor values versus LOP after using LK-CNNC, LK-CNNR, and SK-CNNC algorithms. The measured LOP ranges from -4 dBm to 5 dBm. When the LOP is 0 dBm, the Q-factor of LK-CNNC is 1.58 dB higher than SK-CNNC, and at 1 dBm, the Q-factor of LK-CNNR is 1.47 dB higher than SK-CNNC. It shows that a large kernel is more suitable than a small kernel for extracting information from feature maps constructed using perturbation theory, and higher performance is get. The large kernel designment aided CNNs are proven to be more suitable for optical fiber communication signal processing than small kernel aided CNNs. Fig. 5 also shows that the performance of classifiers is better than that of regressors, when the LOP is 0 dBm, the 0.53 dB Q-factor improvement is obtained with the proposed LK-CNNC nonlinear equalizer compared with LK-CNNR, and the 0.47 dB Q-factor improvement is obtained with SK-CNNC nonlinear equalizer compared with SK-CNNR, which proves that classifiers is better than regressors for same pattern recognition tasks.

Then we analyze the time complexity of these four equalizers when the system performance remains consistent. The time complexity can be measured by FLOPs of two aspects, the convonlutional layer and the fully connected layer. The required time complexity of LK-CNNC, LK-CNNR, SK-CNNC and SK-CNNR are 93925, 101760, 129416, 141955, respectively. From above results, it can be calculated that the time complexity of LK-CNNC reduces 27.4% compares to SK-CNNC, and the time complexity of LK-CNNR reduces 28.3% compares to SK-CNNR. It proves that the large kernel design can effectively reduces computing costs, and classifiers require less computation than regressors.

V. Conclusion

In this paper, the effectiveness of the proposed LK-CNNC nonlinear equalizer is experimentally demonstrated on 120-Gb/s PDM 64-QAM coherent optical communication system transmitted over 375 km distance. The results show that large kernel designment can effectively excavate characteristic of feature maps constructed by perturbation theory because of its large ERF. And when LOP is 0 dBm, the ∼1.58dB Q-factor improvement is obtained with the proposed LK-CNNC nonlinear equalizer compared with SK-CNNC, and when the system performance remains consistent, time complexity is reduced by about 27.4%.

Acknowledgment

This work is supported by the National Key Research and Development Program of China (2021YFB2900703), and the National Natural Science Foundation of China (62075014).

Disclosures

The authors declare no conflicts of interest.

Data availability

Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

[1] Zhou J., Zhang J., Zhao X., et al. “Advanced nonlinearity equalizer with TC-NL-MLSE for transmitting beyond 200-Gb/s PAM-8 in IM/DD systems,” J. Optical Express, vol. 30, no. 21, pp. 37416-37425, 2022.

[2] Wang L., Gao M., Zhang Y., et al. “Optical Phase Conjugation with Complex-Valued Deep Neural Network for WDM 64-QAM Coherent Optical Syetems,” J. IEEE Photonics Journal, 2021.

[3] A. Amari, O. Dobre, R.Venkatesan, O. Kumar, P. Ciblat and Y. Jaouen, “A survey on fiber nonlinearity compensation for 400 Gb/s and beyond optical communication systems,” J. IEEE Communications Surveys and Tutorials, vol. 19, no. 4, pp. 3097-3113, 2017.

[4] E. Ip and J. Kahn, ”Compensation of dispersion and nonlinear impairments using digital backpropagation,” J. Journal of Lightwave Technology, vol. 26, no. 20, pp. 3416-3425, 2008.

[5] Ding J., Liu T., Xu T., Hu W., Popov S., Mark S. Leeson, Zhao J., and Xu T. H., ” Intra-channel Nonlinearity Mitigation in Optical Fiber Transmission Systems Using Perturbation-based Neural Network, ” J. Journal of Lightwave Technology, vol. 40, no. 21, pp. 7106-7116, 2022.

[6] S. Zhang, F. Yaman, K. Nakamura, et al., ”Field and lab experimental demonstration of nonlinear impairment compensation using neural networks,” Nature Communications, vol. 10, no. 3033, Jul.10, 2019.

[7] Q. Fan, G. Zhou, T. Gui, et al., ”Advancing theoretical understanding and practical performance of signal processing for nonlinear optical communications through machine learning,” Nat. Commun., vol. 11, no. 3694, Jul. 23, 2020.

[8] C. Li, Y. wang, L. Han, et al., ”Convolutional Neural Network-Aided DP-64 QAM Coherent Optical Communication Systems,” IEEE Journal of Lightwave Technology, vol. 40, no. 9, 2022.

[9] D. Xiaohan, Z. Xiangyu and Jungong Han and Guiguang Ding, “Scaling Up Your Kernels to 31×31: Revisiting Large Kernel Design in CNNs,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2022.

[10] C. Peng, X. Zhang, G. Yu, G. Luo and J. Sun, “Large Kernel Matters — Improve Semantic Segmentation by Global Convolutional Network,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1743-1751, July, 2017.

[11] Liu S., Chen T., Chen X., Chen X., Xiao Q., and Wu B. “Scaling UP Kernels beyond 51 × 51 Using Sparsity,” IEEE Conference on International Conference on Learning Representations (ICLR), 2023.