Investigating the Performance and Suitability of Neural Network Architectures for Nonlinearity Mitigation of Optical Signals

Vegenshanti Dsilva[1], Isaac Sackey[1], Gregor Ronniger[1], Guillermo von Hünefeld[1,2], Binoy Chacko[1], Colja Schubert[1], and Ronald Freund[1,2]

[1] Fraunhofer Institute for Telecommunications, HHI, Einsteinufer 37, 10587 Berlin, Germany (Email: vegenshanti.valerian@hhi.fraunhofer.de)
[2] Technische Universität Berlin, Fachgebiet Photonische Kommunikationssysteme, Einsteinufer 25, 10587 Berlin, Germany

Abstract *We compare three different neural network architectures for nonlinearity mitigation of 32 GBd OOK and QPSK signals after transmission over a dispersion-compensated link of 10-km SSMF and 10-km DCF. OSNR gains up to 2.2 dB were achieved using reservoir networks, suitable for fast training. ©2022 The Author(s)*

Introduction

Current optical networks are moving towards new digital services (e.g. low-latency high-definition streaming media and cloud computing) that will revolutionize the 5G and beyond infrastructures. One key underpinning approach driving this goal is brain-inspired machine-learning (ML) assisted concept which has been shown to improve computational speed and latency in comparison to architectures based on von Neumann [1-3].

Another critical aspect required to achieve such high-end services relies on high-capacity data transmission. However, moving from lower spectrally efficient modulation formats to higher-order (multi-level) modulation formats, as a means to increase capacity, requires higher received optical signal-to-noise ratios (OSNR). The use of higher launch powers to realize a high OSNR introduces fiber Kerr nonlinear distortions on the transmitted data signals. Until recently, classical approaches in both optical [4-7] and digital domains [8-10] have been used for nonlinearity mitigation of telecom signals. However, leveraging the benefits of ML-based artificial neural networks (NN) for nonlinearity mitigation is a promising paradigm shift for improving signal quality with high computational speed and low latency [11-14].

In this contribution, we employ numerical simulations to investigate the nonlinearity mitigation performances of three key artificial NNs (i.e., reservoir computing network (RCN), recurrent neural network (RNN) and feedforward neural network (FNN)). After independently transmitting two different modulation formats (32 GBd single-polarization non-return-to-zero (NRZ) on-off keying (OOK) data signal and 32 GBd quadrature phase-shift keying (QPSK) data signal) over a cascade of a 10-km standard single-mode fiber (SSMF) and a 10-km dispersion-compensated fiber (DCF), we compare and report on the nonlinear mitigation performances of the investigated NNs for those modulation formats.

Data Generation and Simulation Setup

To model the various NN architectures and evaluate their nonlinear mitigation performances, we use a numerical simulation based on VPIphotonics Design Suite 11.1 for generating datasets with different modulation formats and noise levels. The whole simulation setup is depicted in Fig. 1. The datasets are generated for the two investigated modulation formats: 32 GBd NRZ OOK and 32 GBd QPSK. To generate the optical data signals, pseudo-random binary sequences (PRBS) of order 2^{16} are pulse-shaped, reconstructed by a digital-to-analog converter (DAC) and applied to an IQ modulator to modulate a continuous wave emitted by an

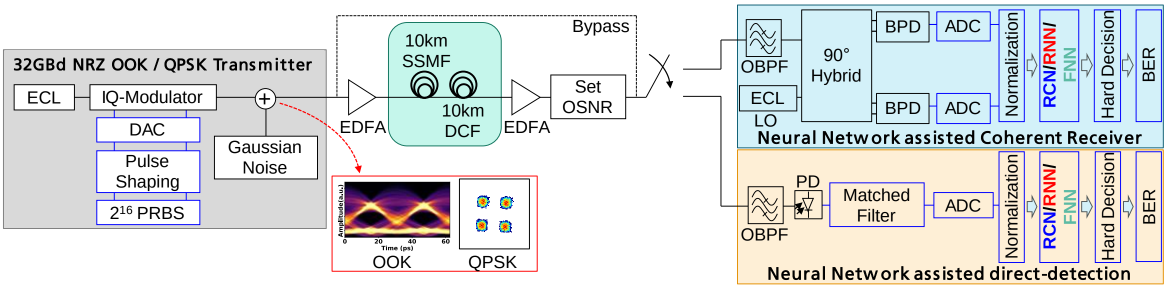

Fig. 1: Simulation setup depicting a transmitter capable of generating 32 GBd NRZ OOK or 32 GBd QPSK signals, a transmission link, and corresponding neural network based direct-detection receiver and a coherent receiver for the signals.

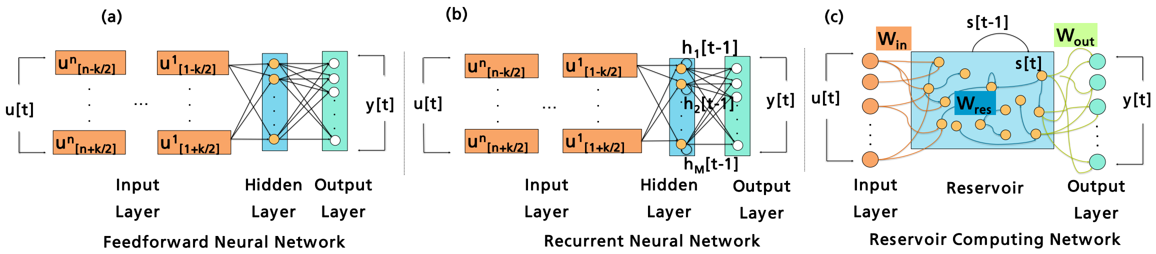

Fig. 2: Architectures depicting the structures of the input, hidden and output layers of (a) FNN, (b) RNN and (c) RCN. The input signal and output mitigated signal are denoted by *u* and *y*, respectively of length *n*. The total number of past and future samples taken from the distorted signal in the input vector at each time instant *t* is denoted by *k*.

external cavity laser (ECL, 1552.52 nm, +10 dBm). We intentionally introduced an additive Gaussian noise to the generated data with the aim of increasing the interplay of amplitude noise and nonlinearity in the transmission fiber. The generated data is received after it is transmitted over a cascade of a 10-km SSMF and a 10-km DCF. The Erbium-doped fiber amplifier (EDFA, 5.5-dB noise figure) before the link is used to set the launch power of the data to +13 dBm. Note that the launch power is kept fixed in all of the investigated cases. The primary focus of our work is on nonlinearity mitigation, thus we use a DCF, after the SSMF, to minimize the accumulated dispersion of the received data. After compensating for the transmission losses using another EDFA (5.5-dB noise figure), a set OSNR stage is used to emulate a further degradation on the data signal and allow for OSNR variation. In order to train the NNs, the transmitter output signal bypassing the link is used as a target signal in the training and also for visualization. At the receiver, we employ an optical bandpass filter (OBPF) to suppress the out-of-band amplified spontaneous emission noise. The OOK signal is detected using a single-ended photodiode (PD) followed by a matched filter to optimize the received SNR. However, a single-polarization coherent receiver is used to receive the QPSK signal. The photocurrents are sampled at 8 samples per symbol for OOK, 4 samples per symbol for QPSK and normalized to have the same root-mean-square amplitude as the target before feeding the data to the NNs. The NN output is then resampled using optimal thresholds and sample phases and decoded to generate the output bitstream. Comparison with the transmitted bitstream yields the BER.

Feedforward Neural Network (FNN)

Fig. 2(a) depicts the architecture of a time-delayed FNN. It is implemented using *TensorFlow* library [15]. In our realization, we built the input vector u^t to the model using a shifting time window (*k+1* samples length) to contain *k/2* samples from both, the past and the future, to predict a single output sample. For time instants with insufficient past and future samples, the remainder values are set to zero. Note that FNN does not store sequential information of the input [16]. Based on the performance and training time, a window of 5 symbols is chosen as a reasonable size. Thus, the parameter k is equal to 40 and 20 for OOK and QPSK signals, respectively. After model selection, the hidden layer consists of 64 neurons, rectified linear unit *(ReLU*) activation function for model trained on OOK and 32 neurons, exponential linear unit (*ELU*) activation function for model trained on QPSK. The output layer provides the desired mitigated signal using a linear transfer function.

Recurrent Neural Networks (RNN)

In its simplest form, the RNN architecture contains recurrent connections over its *M* hidden layer nodes which enables it to capture temporal dynamics of a signal [16] as shown in Fig. 2(b). Similar to FNN architecture, the input vector to RNN is also time-delayed. The model trained on OOK contains 32 recurrent units with *ReLU* activation while the model trained on QPSK contains 128 recurrent units with hyperbolic tangent (*tanh*) activation in the hidden layer. The hidden layer is followed by the linear output layer predicting the desired equalized signal. The implementation is in *TensorFlow*.

Reservoir Computing Network (RCN)

Fig. 2(c) depicts the RCN architecture [17]. The distorted signal, *u[t],* is sent into the reservoir via input weight matrix W_{in}. The reservoir weight matrix, W_{res}, and W_{in} are used to compute the reservoir state *s[t]*. The hidden layer (reservoir) has a recurrent connection over itself and the reservoir states are used to update the output via the output weights W_{out}. The RCN only trains the output layer using ridge regression. It is implemented using the library *easyesn* [18]. For OOK it has 100 reservoir nodes and spectral radius of 0.45 [17] while for QPSK it uses 300 reservoir nodes and spectral radius of 0.85.

Results and Discussion

The performance metric for our model selection was based on bit-error ratio (BER). We converted

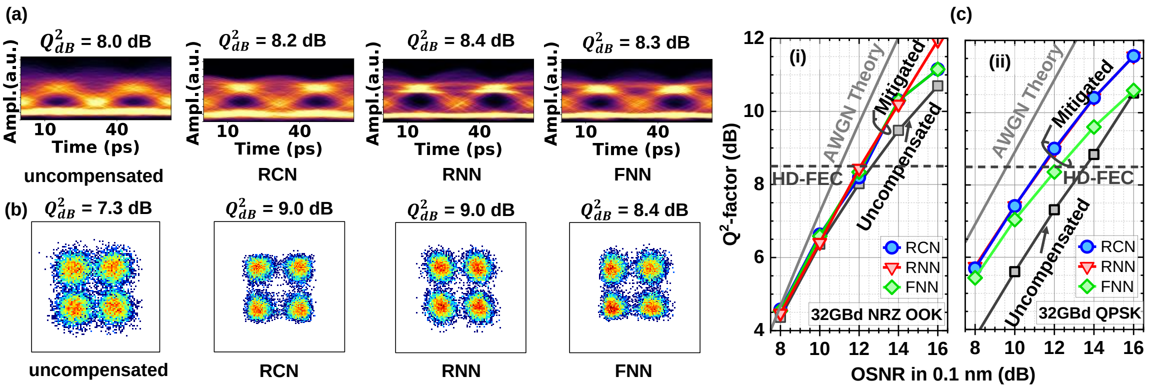

Fig. 3: Nonlinearity mitigation comparison using reservoir computing network (RCN), recurrent neural network (RNN) and feedforward neural network (FNN): (a) Eye diagrams for the uncompensated and mitigated 32 GBd NRZ single-polarization OOK signal at 12-dB OSNR. (b) Constellation diagrams for the uncompensated and mitigated 32 GBd single-polarization QPSK signal at 12-dB OSNR. (c) Q^2-factor vs OSNR performance comparison for the OOK and QPSK signals.

the BER values to Q^2-factors using the relation: $Q^2_{dB} = 20 \times log_{10}\left[\sqrt{2}\, erfc^{-1}(2 \times BER)\right]$. The model selection for FNN and RNN was performed by varying the number of units in the hidden layers (from 4 to 512), activation function (*ReLU, ELU, tanh, sigmoid*), optimizer (*Adam, Stochastic Gradient Descent, Root Mean Squared Propagation)* and learning rate *(*10^{-3}*,* 10^{-2}*,* 10^{-1}*).* Increasing the hidden layer units in FNN up to a certain extent improved the performance but further increase led to saturation. The number of hidden layer units for RNN do not follow a specific pattern and hence the best performing RNN was chosen after varying the hidden units. The number of nodes and spectral radii are two important hyper-parameters characterising the memory of a RCN. Increasing the reservoir size improves the Q^2-factor until it eventually stays constant between 300 and 500 nodes. However, increasing the size of the hidden layer requires longer training time. The overall training for RCN was ~ 5 times faster than for RNN.

We exemplarily show comparison of eye diagrams for OOK in Fig. 3(a) and constellation diagrams for QPSK in Fig 3(b), respectively, for the different NN models at 12-dB OSNR. They indicate that all three models are able to mitigate distortion to a great extent. Fig. 3(c) summarises the Q^2-factor vs OSNR results for OOK (see Fig. 3(c)(i)) and QPSK (see Fig. 3(c)(ii)) signals. While for OOK the models give moderate OSNR gains of 0.3 dB (RCN), 0.4 dB (FNN) and 0.5 dB (RNN), the OSNR gains using a QPSK signal is more pronounced with 1.4 dB for FNN and 2.2 dB for both RNN and RCN. All gains were measured at the hard-decision forward-error-correction (HD-FEC) threshold (i.e., $Q^2_{dB} = 8.5\, dB$).

One critical difference is that the RCN model uses 1 sample per time instant to perform signal equalisation while FNN and RNN are fed with 41 samples (OOK) and 21 samples (QPSK) each time instant to achieve a comparable performance to the RCN. The complexity of training increases as the FNN and RNN models get deeper and wider. Both FNN and RNN have significant limitations in terms of complex input vector and longer training times. Owing to these limitations, RCN is a favourable candidate to utilise the potential of machine learning for signal equalization in a practical scenario.

Conclusion

In this work, we compared different approaches, based on neural networks to mitigate typical nonlinear impairments in optical transmission links. We focused on OOK and QPSK, but plan to extend our approach to higher-order modulation formats and more challenging link conditions. OSNR gains of up to 2.2 dB were realized using our networks. While RCNs are not in widespread use for nonlinearity mitigation, they achieved similar gains compared to RNNs, which are used more often in this application. At the same time RCNs allowed for a 5-fold reduction in training time, which becomes critical, when a more dynamic training during link operation is envisaged. So far, we have performed signal equalization using a supervised learning approach with a consistent training dataset to learn model weights for prediction. The emerging area of semi-supervised learning [19, 20] can be explored as a future scope to develop an adaptive learning equalizer that further decreases the data collection and training efforts.

Acknowledgements

Funded by the German Bundesministerium für Bildung und Forschung (BMBF) under the 6G-RIC project with grants 16KISK020K, 16KISK030, and under the CELTIC-NEXT AI-NET-PROTECT project with grant 16KIS1282.

References

[1] A. S. Cassidy, R. Alvarez-Icaza, F. Akopyan, J. Sawada, J. V. Arthur, P. A. Merolla, P. Datta, M. G. Tallada, B. Taba, A. Andreopoulos, A. Amir, S. K. Esser, J. Kusnitz, R. Appuswamy, C. Haymes, B. Brezzo, R. Moussalli, R. Bellofatto, C. Baks, M. Mastro, K. Schleupen, C. E. Cox, K. Inoue, S. Millman, N. Imam, E. McQuinn, Y. Y. Nakamura, I. Vo, , C. Guo, D. Nguyen, S. Lekuch, S. Asaad, D. Friedman, B. L. Jackson, M. D. Flickner, W. P. Risk, R. Manohar and D. S. Modha, ''Real-time Scalable Cortical Computing at 46 Giga-Synaptic OPS/Watt with ~100× Speedup in Time-to-Solution and ~100,000× Reduction in Energy-to-Solution, '' International Conference for High Performance Computing, Networking, Storage and Analysis, vol. 2015-January, pp. 27-38, IEEE Computer Society, 2014, DOI: https://doi.org/10.1109/SC.2014.8

[2] M. Davies, N. Srinivasa, T. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, Y. Liao, C. Lin, A. Lines, R. Liu, D. Mathaikutty, S. McCoy, A. Paul, J. Tse, G. Venkataramanan, Y. Weng, A. Wild, Y. Yang and H. Wang, ''Loihi: A Neuromorphic Manycore Processor with On-Chip Learning,'' IEEE Micro, vol. 38, no. 1, pp. 82-99, 2018, DOI: https://doi.org/10.1109/MM.2018.112130359

[3] J. Zhu, and P. Sutton, ''FPGA Implementations of neural networks - a survey of a decade of progress,'' Proc. International conference on field programmable logic and applications, pp. 1062–1066, Springer, 2003, DOI: https://doi.org/10.1007/978-3-540-45234-8_120

[4] I. Sackey, T. Richter, M. Nölle, M. Jazayerifar, K. Petermann, J. K. Fischer, and C. Schubert, ''Qualitative Comparison of Kerr Nonlinearity Mitigation Schemes in a Dispersion-Managed Link for 4 × 28-GBd 16-QAM Signals,'' J. Lightwave Technol. vol. 33, pp. 4815-4825, 2015, DOI: https://doi.org/10.1109/JLT.2015.2489703

[5] S. L. I. Olsson, B. Corcoran, C. Lundström, T. A. Eriksson, M. Karlsson, and P. A. Andrekson, ''Phase sensitive amplified transmission links for improved sensitivity and nonlinearity tolerance,'' J. Lightwave Technol., vol. 33, no. 3, pp. 710-721, 2015. DOI: https://doi.org/10.1109/JLT.2014.2367096

[6] T. Richter, R. Elschner, and C. Schubert, "QAM phase-regeneration in a phase-sensitive fiber-amplifier,'' IET Conference Publications vol. 2013, pp. 489–491, 2013, DOI: https://doi.org/10.1049/cp.2013.1436

[7] F. Da Ros, D. Vukovic, A. Gajda, K. Dalgaard, L. Zimmermann, B. Tillack, M. Galili, K. Petermann, and C. Peucheret, ''Phase regeneration of DPSK signals in a silicon waveguide with reverse-biased p-i-n junction,'' Optics Express, vol. 22, no. 5, pp. 5029, 2014, DOI: https://doi.org/10.1364/oe.22.005029

[8] E. Ip, ''Nonlinear Compensation Using Backpropagation for Polarization-Multiplexed Transmission,'' J. Lightwave Technol., vol. 28, no. 6, pp. 939-951, 2010, DOI: https://doi.org/10.1109/JLT.2010.2040135

[9] R. Weixer, J. Koch, P. Plany, S. Ohlendorf, and S. Pachnicke, ''Mitigation of nonlinear impairments by using support vector machine and nonlinear volterra equalizer,'' Applied Sciences (Switzerland), vol. 9, no. 18, 2019, DOI: https://doi.org/10.3390/app9183800

[10] Q. Fan, G. Zhou, T. Gui, C. Lu, and A. P. T. Lau, ''Advancing theoretical understanding and practical performance of signal processing for nonlinear optical communications through machine learning,'' Nature Communications, vol. 11, no. 1, 2020, DOI: https://doi.org/10.1038/s41467-020-17516-7

[11] S. M. Ranzini, R. Dischler, F. Da Ros, H. Bulow, and D. Zibar, ''Experimental Demonstration of Optoelectronic Equalization for Short-reach Transmission with Reservoir Computing,'' 2020 European Conference on Optical Communications, ECOC 2020. Institute of Electrical and Electronics Engineers Inc., 2020, DOI: https://doi.org/10.1109/ECOC48923.2020.9333372

[12] M. Schaedler, C. Bluemm, M. Kuschnerov, F. Pittalà, S. Calabrò, and S. Pachnicke, ''Deep neural network equalization for optical short reach communication,'' Applied Sciences (Switzerland), vol. 9, no. 21, 2019, DOI: https://doi.org/10.3390/app9214675

[13] J. C. Gomez Diaz, H. Zhao, Y. Zhu, S. Palermo, and S. Hoyos, ''Recurrent Neural Network Equalization for Wireline Communication Systems,'' IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 69, no. 4, pp. 2116-2120, April 2022, DOI: https://doi.org/10.1109/TCSII.2022.3152051.

[14] O. Kotlyar, M. Kamalian-Kopae, M. Pankratova, A. Vasylchenkova, J. E. Prilepsky, and S. K. Turitsyn, ''Convolutional long short-term memory neural network equalizer for nonlinear Fourier transform-based optical transmission systems,'' Optics Express, vol. 29, no. 7, pp. 11254, 2021, DOI: https://doi.org/10.1364/oe.419314

[15] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, ''TensorFlow: A system for large-scale machine learning,'' Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016, pp. 265–283, USENIX Association, 2016, DOI: https://doi.org/10.48550/arXiv.1605.08695

[16] I. Goodfellow, Y. Bengio, and A. Courville, ''Deep Learning,'' MIT Press, 2016

[17] M. Lukoševičius, ''A practical guide to applying echo state networks,'' Lecture Notes in Computer Science (Including Subseries Lecture Notes, Artificial Intelligence and Lecture Notes in Bioinformatics), 7700 LECTURE NO, pp. 659–686, 2012, DOI: https://doi.org/10.1007/978-3-642-35289-8_36

[18] R. Zimmermann, and L. Thiede, https://pypi.org/project/easyesn/0.1.4.5/ ,7 April 2018.

[19] L. Beyer, X. Zhai, X. Oliver, and A. Kolesnikov, ''S4L: Self-supervised semi-supervised learning,'' Proceedings of the IEEE International Conference on Computer Vision 2019, pp. 1476–1485, Institute of Electrical and Electronics Engineers Inc., 2019, DOI: https://doi.org/10.1109/ICCV.2019.00156

[20] L. Jing, and Y. Tian, ''Self-Supervised Visual Feature Learning with Deep Neural Networks: A Survey,'' IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Computer Society, 2021, DOI: https://doi.org/10.1109/TPAMI.2020.2992393