Convolutional Recurrent Machine Learning for OSNR and Launch Power Estimation: A Critical Assessment

Hyung Joon Cho, Siddharth Varughese, Daniel Lippiatt, and Stephen E. Ralph
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
Author e-mail address: stephen.ralph@ece.gatech.edu

Abstract: Using waveforms from three distinct stages of signal demodulation, we assess performance, computational efficiency, and benefits of using convolutional recurrent neural networks to simultaneously and independently estimate OSNR and launch power within a multi-channel system.

1. Introduction

Neural networks have recently received considerable attention and provide estimations of various optical performance monitoring (OPM) signal metrics including optical signal-to-noise ratio (OSNR), quality of transmission, and bit-error-ratio. However, the benefits of and sometimes the need for machine learning (ML) techniques are not clear even when estimations are made with good accuracy. Some efforts appear not to require ML or to overfit the ML tools and some efforts may have inadvertently trained on improper proxies of the desired parameter. ML is best used when either a physical model or input variables are not available or when computational efficiency can be substantially improved. For example, a convolutional neural network (CNN) that operates exclusively on a constellation image has been shown to accurately estimate OSNR [1]. Certainly, in DWDM systems, OSNR is a primary link metric yet it is challenging to assess using conventional spectral techniques due to tight channel spacing. Some methods including recurrent neural network (RNN) that operate on a sequence and extract features based on information from previous sequences resort to preprocessing sequences for OPM [2]. Furthermore, the generalized OSNR is impacted by both ASE and nonlinear contributions. The latter may be primarily in-band and only available when symbol power is non-zero. Since most systems operate with a balance between amplified spontaneous emission (ASE) limited OSNR and nonlinear noise, efforts have to be made to assess and separate these noise sources.

The amount and characteristics of the limiting noise are affected at various stages of demodulation. Furthermore, it may be beneficial to quantify the noise independently of the demodulation since this allows real time optimization of the demodulation methods, assessment of the demodulation process, and may also provide insight into network health. Thus, assessment of linear and nonlinear noise at various stages of demodulation is beneficial. We choose three distinct points within the demodulation process to assess OSNR and launch power (LP); 1) immediately after digitization, 2) after chromatic dispersion (CD) and polarization mode dispersion (PMD) correction, and 3) after the final stage of equalization, Fig. 1. At each point, we critically assess the performance, computational complexity, and benefits of using ML, specifically we implement convolutional recurrent neural networks (CRNN) which are a hybrid structure of CNN and RNN. We assess the simultaneous and independent estimation performance of OSNR and LP using CRNN on 32 GBaud dual-polarized (DP) QPSK signals with no extra preprocessing of the waveforms.

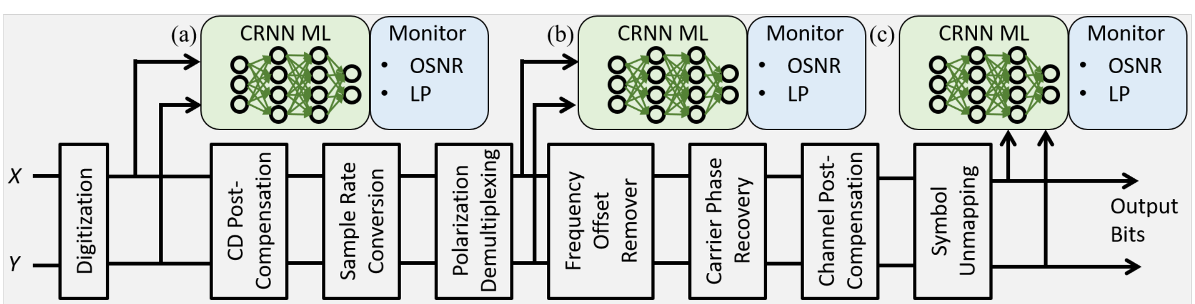

Fig. 1. Demodulation scheme showing three distinct stages where waveforms are captured to estimate OSNR and LP, (a) immediately after digitization, (b) after CD and PMD correction, and (c) after the final stage of equalization

2. Convolution Recurrent Neural Networks for Optical Performance Monitoring

CRNN is a hybrid architecture that utilizes the features of CNN and RNN to extract noise features from the waveform. CNN extracts features through several convolutional stages that convolve the input with various one-dimensional filters to enhance image features such as edges and shapes. The pooling layer efficiently reduce the computational

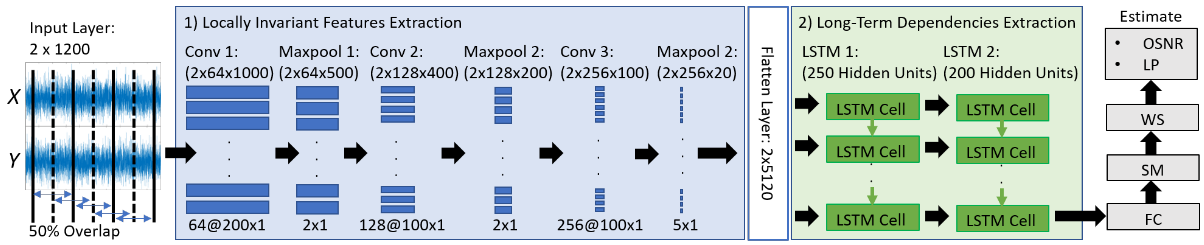

Fig. 2. CRNN Structure for Estimation OSNR and Launch Power. There is dropout layer of 0.1 between each CNN and LSTM layers. All of input weight initializers are Glorot normal initializers and all CNN activation functions are leaky ReLU of 0.1.

complexity by reducing the by employing a one-dimensional sliding window with a specific stride and choosing the maximum value in every window. While CNN is a strong tool in capturing the spatial features, it considers only the current input. Therefore, it may not be sufficient when estimating noise characteristics that include colored noise with significant correlations. On the other hand, RNN can capture temporal dynamic behaviors within a sequence owing to its recurrent architecture. Long short-term memory (LSTM), a special type of RNN, captures both short and long-term dependencies of a sequence by using its internal memory along with a forget gate that calculates the proportion of the useful data to be kept in each of the LSTM cell [3,4]. However, LSTM is more challenging to train and may be more vulnerable to overfitting than CNN [5].

Figure 2 shows the architecture employed in this work. The hybrid structure complements their individual limitations while maximizing their advantages. The 1D-convolutional layers extract the mid-level and locally invariant features from the noises of the input waveforms while the LSTM cells capture the long and short-range contextual information from those features. The sliding window technique is applied so that successive waveforms are captured with some overlap to enhance noise features. We used 5 windows of only 1200 symbols with 50% overlaps from the PRBS-15 waveforms captured at random moments. The features learned in the last LSTM cell are post-processed through fully-connection (FC) to softmax (SM) layers and weighted sums (WS) for the classification of OSNR and LP. We captured 175 waveforms for each OSNR ranging from 11 dB to 17 dB with 1 dB increment and LP ranging from 3 dBm to 9 dBm with 1 dBm increment. The OSNR range chosen is appropriate for QPSK. However, our CRNN tool will work with higher OSNR values since we demonstrate that our ML tool can learn the noise features in all the unique combinations of OSNR and LP. We allocated 80% of the data sets for training and 20% for testing. The numbers of kernels, layers, and strides are hyperparameters and optimally tuned with the waveforms captured from R-Soft OptSim® prior to our experiments.

3. Experimental Setups and Results

Figure 3(a) shows our optical testbed consisting of three DWDM channels each with DP 32-GBaud QPSK signals. The signals are transmitted across 9 spans of 90 km SSMF. ASE noise is loaded before a Wavelength Selective Switch (WSS) that limits bandwidth to 37.5 GHz. Conventional DSP is used for processing the data [6].

3.1. Immediately After Digitization

Figure 3(b) shows the OSNR estimation errors from waveforms captured immediately after digitization where maximum information regarding noise characteristics should be available. The mean OSNR estimation error is ~1 dB with a maximum error of 2 dB. Figure 3(c) shows the estimated LP values for all values of OSNR. The LP estimation improves with launch power since the nonlinearity signature is greater. This observation was consistent for all 49 configurations and is common to all stages of information extraction. This demonstrates the challenges of estimating launch power when the signal is ASE white noise limited. It is possible that the ML at this stage is also limited by the impact of CD. Furthermore, it may benefit if the testing is accomplished with noise added in a distributed fashion to enhance the nonlinear features. We note again that we do not using any preprocessing of the digitized signal and that subsequent demodulation is not required. The potential advantage of extracting information without conventional DSP is likely limited due to the high computational efforts needed to process the raw digitized waveform.

3.2. After CD and PMD Correction

Figure 3(d) shows the OSNR estimation errors from waveforms captured after CD and PMD correction. Here we observe the best accuracy of the three cases with a maximum error < 1 dB. The overall average error was 0.23 dB. The launch power estimations errors were also very good with < 1 dBm deviation, Fig. 3(e). The DSP stages after the digitization refined the waveforms and enables ML to learn features of noise variations for all 49 configurations. The CD and PMD corrections provided much waveforms where symbols are more readily extracted from the noise-like

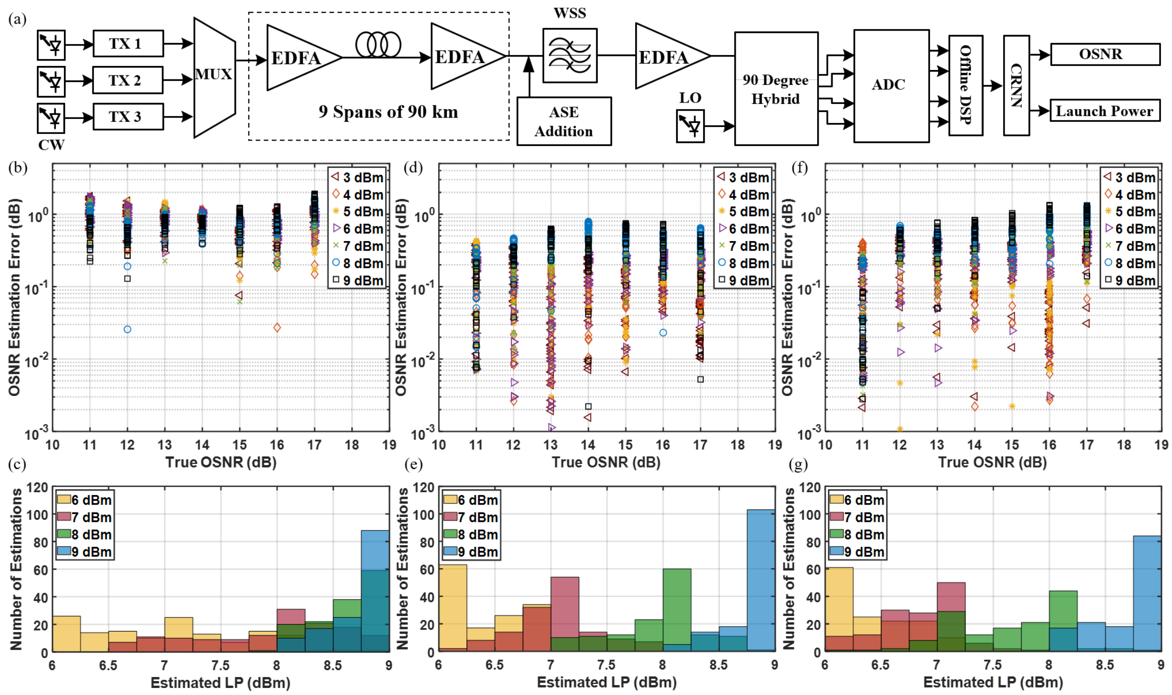

Fig. 3. (a) Experiment testbed for DP 32-GBaud QPSK signals. Immediately after digitization: (b) OSNR errors (c) Estimated LP. After CD compensation and PDM: (d) OSNR errors (e) Estimated LP. After the final stage of equalization: (f) OSNR errors (g) Estimated LP.

features, thereby enabling the ML to learn how both ASE and nonlinear noise varies from the different configurations of OSNR and LP more effectively and accurately. Even though the signal has gone through some of the DSP steps, it is still beneficial to extract signal parameters at this point since the accuracy is higher and the estimates can be used to improve the remainder of the DSP chain. Additional signal preprocessing may further reduce the ML computational burden and result in better estimations. Indeed, this may be an optimum point to extract noise information. Lastly, in the event of poor signal metrics, the demodulation can also be aborted saving time and energy.

3.3. After Final Stage of Equalization

Figure 3(f) shows the OSNR estimation errors from waveforms after full demodulation. The accuracy is slightly worse than the previous case with a maximum error of 1.3 dB. The overall average error was 0.36 dB. Similarly, the LP estimation shown in Fig. 3(g) is slightly worse than the previous case. At this stage, the demodulation scheme has attempted to compensated for all impairments and noise features are either distorted, minimized or removed making it harder for the ML algorithm to detect these features. Moreover, the algorithm at higher OSNRs demonstrate overfitting since the noise features are less discernible and the algorithm attempts to learn the transmitted pattern itself. Even though data extraction may be easier at this point in the demodulation, accuracies are compromised due to the equalization schemes. It may be easier and faster to analyze the constellation diagrams or BERs.

4. Conclusion

We presented a CRNN-based ML technique that simultaneously and independently estimates OSNR and LP using raw waveforms captured at three distinct stages during the demodulation process without any extra preprocessing. The CRNN's ability to learn the noise features for unique configuration of OSNR and LP was experimentally demonstrated. We experimentally validated that OSNR and LP can be estimated with < 1 dB and < 1 dBm errors respectively from the waveforms captured right after the PDM. We note that some efforts may inadvertently train on improper proxies of the desired parameter and taking features at a suitable DSP stage is crucial for accurate OPM.

5. References

[1] Cho et. al., "Convolutional Neural Networks for Optical Performance Monitoring," AVFOP, WD2, Arlington, VA, Nov. 2019.
[2] Wang et. al., "OSNR and Nonlinear Noise Power Estimation for Optical Fiber Communication Systems Using LSTM Based Deep Learning Technique," Opt. Express, vol. 26, no. 16, Aug. 2018, pp. 21346 - 21357.
[3] Jozefowicz et. al., "An Empirical Exploration of Recurrent Network Architectures," ICML, July 2015, pp. 2342–2350.
[4] Greff et. al., "LSTM: A Search Space Odyssey," IEEE Transaction on Neural Networks and Learning Systems, vol. 28, no. 10, Oct. 2017.
[5] Zhu et al., "Multimodal Gesture Recognition Using 3-D Convolution and Convolutional LSTM," IEEE Access, vol. 5, March 2017.
[6] S. Varughese et al., "Frequency dependent ENoB requirements for M-QAM optical links: an analysis using an improved digital to analog converter model," in J. Lightwave Technol. 36, 4082-4089 (2018).