



Experimental validation of CNNs versus FFNNs for time- And energy-efficient EVM estimation in coherent optical systems

Downloaded from: <https://research.chalmers.se>, 2024-04-23 06:33 UTC

Citation for the original published paper (version of record):

Fan, Y., Udalcovs, A., Natalino Da Silva, C. et al (2021). Experimental validation of CNNs versus FFNNs for time- And energy-efficient EVM estimation in coherent optical systems. *Journal of Optical Communications and Networking*, 13(10): E63-E71. <http://dx.doi.org/10.1364/JOCN.423384>

N.B. When citing this work, cite the original published paper.

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

Experimental Validation of CNN vs. FFNN for Time- and Energy-Efficient EVM Estimation in Coherent Optical Systems

YUCHUAN FAN,^{1,2} ALEKSEJS UDALCOVS,² CARLOS NATALINO,³ XIAODAN PANG,^{1,2}
RICHARD SCHATZ,¹ MARIJA FURDEK,³ SERGEI POPOV,¹ OSKARS OZOLINS^{1,2*}

¹ School of Engineering Sciences, KTH Royal Institute of Technology, Isafjordsgatan 22, 164 40 Kista, Sweden

² RISE Research Institutes of Sweden, Isafjordsgatan 22, 164 40 Kista, Sweden

³ Department of Electrical Engineering, Chalmers University of Technology, Chalmersplatsen 4, 412 96 Gothenburg, Sweden

*Corresponding author: oskars.ozolins@ri.se

Received XX Month XXXX; revised XX Month, XXXX; accepted XX Month XXXX; posted XX Month XXXX (Doc. ID XXXXX); published XX Month XXXX

Error vector magnitude (EVM) has proven to be one of the optical performance monitoring (OPM) metrics providing the quantitative estimation of the error statistics. However, the EVM estimation efficiency is not fully exploited in terms of complexity and energy consumption. Therefore, in this article, we explore two deep-learning-based EVM estimation schemes. The first scheme exploits convolutional neural networks (CNNs) to extract EVM information from images of constellation diagram in the In-phase/Quadrature (IQ) complex plane or amplitude histograms (AH). The second scheme relies on feedforward neural networks (FFNN) extracting features from a vectorized representation of AHs. In both cases, we use short sequences of 32 Gbaud m-ary quadrature amplitude modulation (mQAM) signals captured before or after a carrier phase recovery (CPR). The impacts of the sequence length, neural network structure, and dataset representation on the EVM estimation accuracy as well as the model training time are thoroughly studied. Furthermore, we validate the performance of the proposed schemes using the experimental implementation of 28 Gbaud 64QAM signals. We achieve a mean absolute estimation error below 0.15% with short signals consisting of only 100 symbols per IQ cluster. Considering the estimation accuracy, the implementation complexity, and the potential energy savings, the proposed CNN- and FFNN-based schemes can be used to perform time-sensitive and accurate EVM estimation for mQAM signal quality monitoring purposes.

<http://dx.doi.org/10.1364/JOCN.99.099999>

1. INTRODUCTION

Optical transport experiences an exponential data traffic increase. This is driven by the explosive growth of Internet services and cloud applications [1,2]. Coherent optical communications pave a way for addressing the capacity crunch and, therefore, these solutions are deployed in evolving metro and long-haul optical networks. Simultaneously, optical performance monitoring (OPM) is gaining a more important role for the timely management of such networks [3,4]. The commonly monitored parameters are optical power, optical signal to noise ratio (OSNR), bit-error-rate (BER), Q-factor, and error vector magnitude (EVM) [5]. EVM is a well-developed quality measure metric for m-ary quadrature amplitude modulation (mQAM) formats. For an optical channel limited by the additive white Gaussian noise (AWGN), an EVM can be mapped to a BER and SNR [6-8]. Traditionally, EVM calculation needs millions of received symbols [6]. However, such a cumulative process meets an implementation complexity issue in terms of the required time and energy consumption. Thanks to intelligent

estimation schemes [9,10], EVM has been promoted as an OPM metric for fast and accurate signal quality monitoring in optical networks. In the literature, Q-factor [11] and BER [12,13] are chosen for the quality of transmission (QoT) estimation, which are typically performed establishing new lightpaths. Besides, OSNR is widely studied for OPM to measure the actual optical signal itself to estimate signal/link quality [14-20]. Monitoring EVM can extend the functionality of the OPM module, which provides more accurate estimation of the performance compared with the OSNR only.

An ideal OPM scheme is expected to be versatile (considering the diversity of signals in a network), simple, and improving the cost-effectiveness of the network. In this regard, deep learning is widely considered [5,11,13-23]. It has strong capability to extract features from different signal representations, such as time-domain symbol sequence [14,15], phase portrait [16], frequency-domain transformation [17], constellation diagram in the In-phase/Quadrature (IQ) complex plane [9,18,21], and amplitude histogram (AH) [10,19,20]. Recently we have proposed a fast and accurate EVM estimation scheme based on a convolutional neural network (CNN) [9,10]. It uses grayscale images of

IQ [9] or AH [10] of a short sequence for signal quality monitoring purposes. However, the processing of these images with a CNN can be computationally demanding. That may overweight the benefits obtained by operating with short signal sequences. Hence, their representation must be changed so they can be inferred with a computationally less demanding machine learning (ML) model such as feedforward neural networks (FFNNs). Particularly, signal AH can be fed into an FFNN model. Moreover, both the processing time and the energy consumption can be further reduced by skipping some DSP blocks, for instance, a carrier phase recovery (CPR) [10]. The higher a QAM order, the higher a CPR energy consumption [24]. Estimating transmission performance is required for both the intermediate network nodes and receiver. OPM devices located in intermediate network nodes can afford limited complexity due to cost constraints [5,19]. Therefore, it is advantageous to implement OPM schemes that are modulation-format-independent and use as simplified DSP routines as possible, unless the estimation accuracy is degraded.

Compared with our previous works on a CNN-based scheme for EVM estimation [9,10], the main contributions of this paper are: first, it provides experimental validation of the proposed scheme; second, it questions the necessity of CNN-based structures and the CPR module to be a part of the signal quality monitoring functionality; and, third, it performs a study on the required training time and the associated energy consumption for the CNN vs. a newly proposed FFNN scheme in the coherent optical system with QAM orders as high as 256. A newly proposed scheme relies solely on a vectorized representation of AHs obtained from short mQAM signal sequences captured before the CPR module in coherent transceivers. This approach reduces the computational requirements, processing time, and energy consumption for the OPM task. The accuracy of the proposed EVM estimator is evaluated using simulation datasets consisting of 32 Gbaud quadrature phase-shift keying (QPSK, $M = 4$ clusters), 16QAM ($M = 16$ clusters), 64QAM ($M = 64$ clusters) and 256QAM ($M = 256$) signals after 2000 km, 1500 km, 1000 km, 300 km long fiber transmission, respectively. Considering that images may better preserve EVM features, we also explore the performance of the CNN-based scheme operating with images of IQs and AHs. Furthermore, the experimental validation of the proposed schemes is also conducted for 28 Gbaud 64QAM signals. In the experimental setup, these signals are impaired by AWGN, phase noise, and inherent implementation penalty. The results show that our proposed FFNN-based EVM scheme maintains the estimation accuracy. Therefore, it is a good candidate to perform time-sensitive and accurate EVM estimation for mQAM signal quality monitoring purposes.

The rest of the paper is organized as follows. Section 2 outlines the proposed EVM estimation schemes and describes the datasets used in

this study. In Section 3, the performance of the proposed schemes is comprehensively evaluated and compared with the previous EVM estimation schemes. Section 4 analyzes the energy consumption for neural network models. The experimental validation is presented in Section 5. Finally, Section 6 concludes the paper.

2. OPERATING PRINCIPLES

In this section, we detail the operating principles of the proposed schemes, including the simulation setup used for the dataset collection, CNN and FFNN structures, and their hyperparameters.

2.1 Datasets Collection

We use VPItransmissionMaker™ [25] to realize a coherent optical transmission system operating at 32 Gbaud as shown in Fig. 1. In the coherent transmitter, the repeated copies of the generated $2^{15}-1$ pseudorandom bit sequence (PRBS) are Gray-mapped onto 2^{19} symbols in a complex plane. The modulated symbols are then filtered by pulse shaper to generate bandwidth-limited signals. The pulse shaper is a root-raised cosine filter with a 0.15 roll-off factor. The considered modulation formats are QPSK, 16QAM, 64QAM, and 256QAM. The Mach-Zehnder-based in-phase and quadrature modulator (IQM) is used to create the optical signal. Then it is amplified by an erbium-doped fiber amplifier (EDFA) and transmitted over a fiber link. Each span consists of a 100km-long standard single-mode fiber (SSMF) and an inline EDFA. The SSMF has a chromatic dispersion coefficient $D = 16\text{e-}6$ s/m², an attenuation coefficient $\alpha = 0.2$ dB/km, and a nonlinear refractive index $n = 2.6\text{e-}20$ m²/W. The transmission system works in linear region by operating with optimized launch power. In the simulations, we consider the maximum transmission distance for QPSK, 16QAM, 64QAM, 256QAM to be 2000 km, 1500 km, 1000 km, 300 km, respectively. Before the transmission, we set the OSNR to 45dB in the transmitter. We measure the OSNR at 0.1nm resolution after a set of transmission spans using an optical spectrum analyzer (OSA). After a certain transmission distance, we save signal waveforms and use them for the DSP and the dataset accumulation.

To quantify the influence of the CPR processing on the EVM estimation accuracy, we collect short signal datasets before and after this module. For the CNN-based scheme, we generate images of AHs and IQs, whereas a vectorized representation of AHs is used for the FFNN-based scheme. The image datasets are used as a reference to benchmark the performance of the FFNN. Examples of 16QAM signal representations are shown in Fig. 1. Before the CPR processing, the signal is dominantly impaired by the phase noise. Therefore, we observe the IQ symbol rotation in the complex plane and as distortions of the

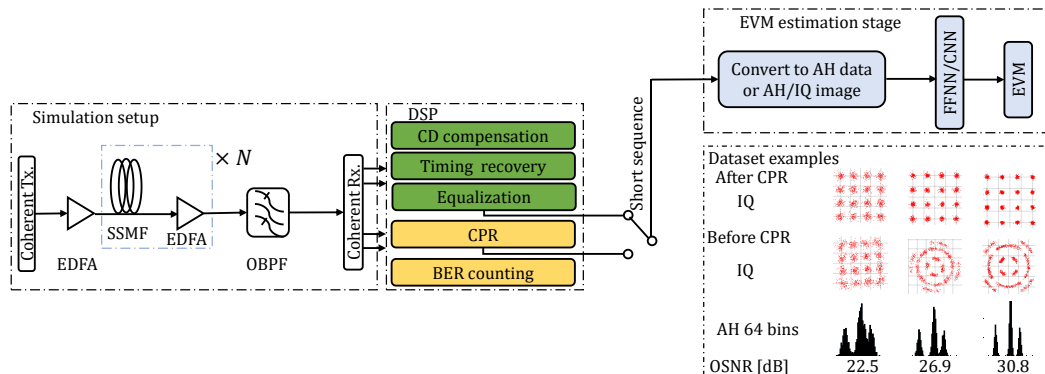


Fig. 1. Schematic diagram of the simulation setup used for the dataset collection. The dataset examples for 16QAM signals with 100 symbols per cluster are shown as insets. EDFA, erbium-doped fiber amplifier; SSMF, standard single-mode fiber; OBPF, optical bandpass filter; DSP, digital signal processing; CD, chromatic dispersion; CPR, carrier phase recovery.

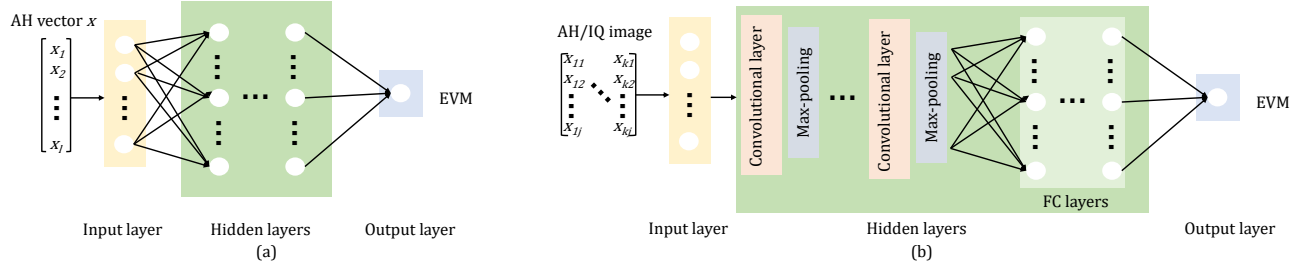


Fig. 2. The deep learning based EVM estimators. (a) FFNN, (b) CNN.

phase-impaired symbol amplitude distributions in the AHs. They display numerical data by grouping them into uniformly spaced amplitude levels called bins [5]. The number of bins can be freely set and used to consider distortions from a low-resolution analog-to-digital converters (ADC). Hence, to assess the impact that the number of bins has on the EVM estimation accuracy, we use AHs with 8, 16, and 64 bins. Finally, to explore how short signal sequences are sufficient for accurate EVM estimation, we generate several dataset options, containing AHs and IQs from $N = 10, 50, 100, 500, 1000$ symbols per cluster in the complex constellation plane (N -symbol/cluster). Every N -symbol/cluster dataset contains 22 simulation scenarios (the product of modulation format and transmission distance (OSNR) options). Each simulation scenario contains 100 samples of vectorized AHs or AHs/IQs images represented by N -symbol/cluster short data sequences. We calculate the true EVM labels from received symbols after the CPR using the k-mean clustering method to obtain centroids for each constellation cluster [26]. Later, each dataset is divided using a 50:25:25 ratio for training, validation, and testing purposes.

2.2 Neural Network Structures

After grooming the accumulated datasets, we construct the corresponding regression models using FFNN and CNN structures for outputting the estimated EVM value. A FFNN, a.k.a. multi-layer perceptron, is an inference model that transforms the input space to the output space through multiple compositions of simple nonlinear functions [27,28]. Our implemented EVM estimator consists of the input layer, hidden layers, and output layer, as depicted in Fig. 2 (a). The AH vector x enters input layer and passes through the constructed FFNN

$$a^m = f_m(W^m \cdot a^{m-1} + b^m), \quad (1)$$

where W^m and b^m denote the weight and bias matrix from layer $m - 1$ to layer m . The $f_m(\cdot)$ is the activation function of layer m . A CNN, however, is composed of convolutional, pooling, and fully connected layers, as shown in Fig. 2 (b). In our case, the input layer receives the 2D array of the AHs/IQs image pixels. Thanks to further convolution operation, the CNN has excellent feature extraction capabilities. A convolutional layer generates feature maps by convolving its input and filters. Each convolutional layer is followed by a pooling layer; we use max-pooling layers to perceive the main features of a certain region. These layers are followed by a fully connected (FC) layer responsible for mapping the feature maps to the output target.

We use a rectified linear unit (ReLU) as an activation function for both types of neural networks to prevent the gradient vanishing problem [29]. An essential work in the regression process is the adjustment of the network weights and biases through the iteration using training data. We randomly initialize all weights and biases in the network and then update them to minimize the loss by backpropagating the gradient of error during the training process [30]. We adopt the mean squared logarithmic error (MSLE) as the loss function to grasp small estimation errors. The MSLE between the true EVM (EVM_t) and the estimated EVM (EVM_e) is defined as:

$$MSLE = \frac{1}{n} \sum_{i=1}^n (\log(EVM_{t_i} + 1) - \log(EVM_{e_i} + 1))^2 \quad (2)$$

where n represents the total number of samples. During the training process, we try to minimize the error; and a larger number of hidden layers allow the neural network to learn more complex decision

Table 1. The performance of the EVM estimator for different neural network structures when operating with 16-bin AHs obtained for the signal sequence of 100-symbol/cluster. MAE are given for a test dataset.

Neural network types	Number of neurons in each layer	MAE [%]	Training time [s] of 200 epochs	FLOPs [G]
FFNN structure	100, 1	3.0	19.8	7.0e-5
	500, 100, 1	0.7	21.4	2.3e-3
	500, 500, 100, 1	0.6	23.0	1.2e-2
	500, 500, 500, 100, 1	0.5	24.8	2.2e-2
	1000, 500, 500, 100, 1	0.4	25.7	3.3e-2
	1000, 500, 500, 500, 100, 1	0.4	30.2	4.3e-2
	2000, 500, 500, 500, 100, 1	0.5	32.9	6.3e-2
The reference CNN structure	4 convolutional layers with 8, 16, 16 and 8 filters; 3-by-3 Kernel; 2 FC layers with 500 and 100 neurons, output layer with 1 neuron	0.5	901.1	5.7

structure. Each neuron in the layer m can receive outputs from neurons in the previous layer $m - 1$ and produce an output propagated to the next layer. This operation can be denoted as

boundaries. However, this may lead to overfitting [31]. Therefore, we optimize the number of hidden layers and neurons in each hidden layer.

3. RESULTS AND DISCUSSION

Table 1 summarizes the performance of the EVM estimator for different FFNN structures defined by the number of neurons in each layer. The performance is compared in terms of mean absolute error (MAE) for the test datasets and training time required for 200 epochs. We use the dataset of 16-bin AHs obtained from the signal sequence of 100 symbols per cluster. The Adam optimization algorithm [32], with the learning rate of $1e-4$, is used to minimize the MSLE loss. The neural network model is implemented using the Keras framework and TensorFlow library [33,34]. Note that the experiments reported in this paper are performed on a 2.4GHz Intel Xeon E5-2630-v3 with 64 GB of random-access memory (RAM) and a GTX TITAN Black graphics processing unit (GPU). This table also includes the corresponding metrics but for a CNN-based scheme that we use as a reference case. The CNN structure is the same as in [9,10]. It consists of 4 convolutional layers (having 3-by-3 kernel size and containing 8, 16, 16, and 8 filters), two fully connected layers with 500 and 100 neurons, and an output layer with one neuron that outputs an estimated EVM value. As input, we use grayscale images of AHs (not vectorized AHs as it is for the FFNN). The results show that the MAE steadily decreases with the number of layers up to a certain number of layers. Increasing further the number of layers in the FFNN

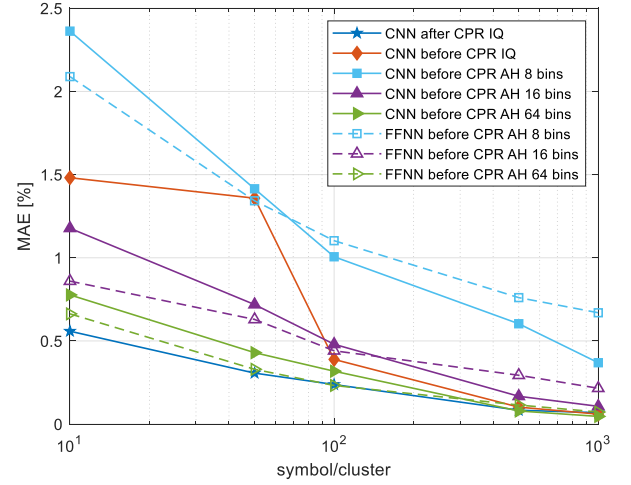


Fig. 3. EVM estimation errors versus N symbols/cluster for various mQAM signal representations in the test datasets.

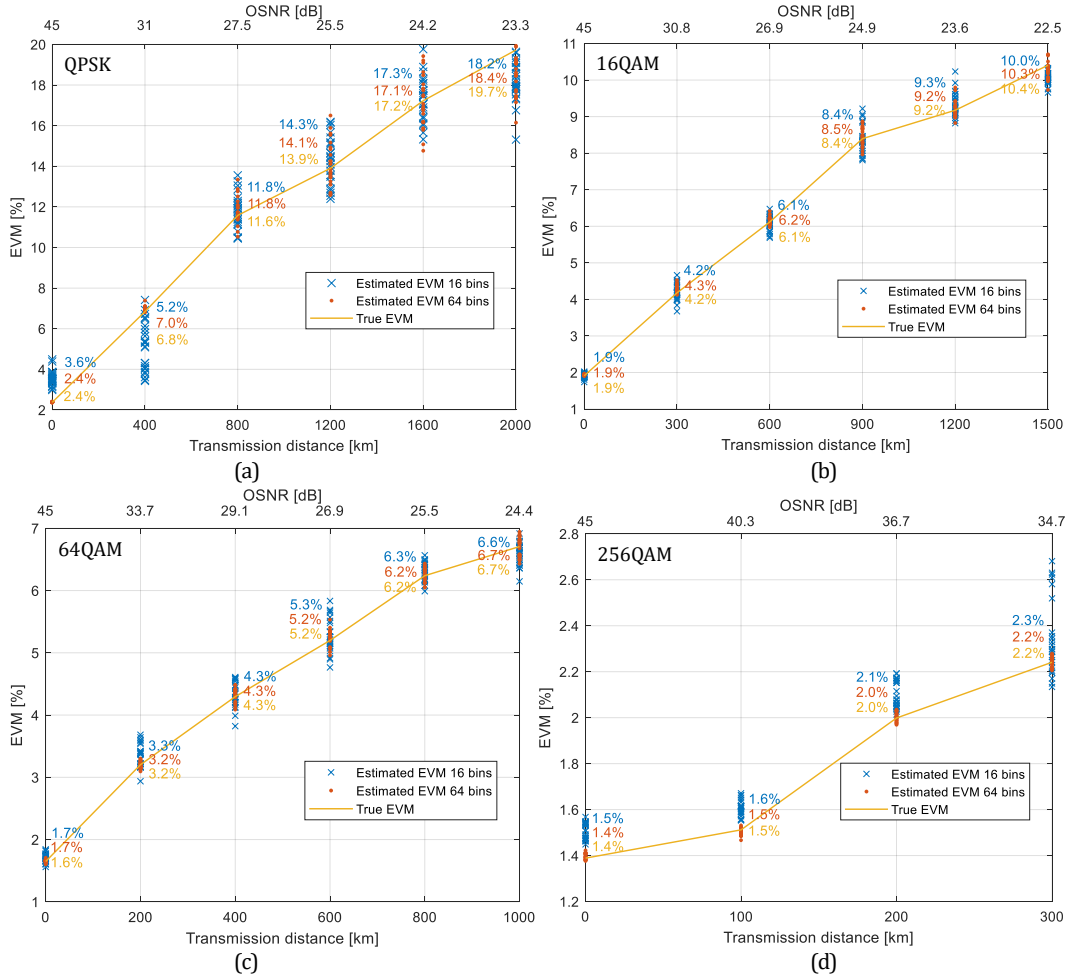


Fig. 4. True EVM versus FFNN-estimated EVM for the test dataset as a function of transmission distance and OSNR for (a) QPSK, (b) 16QAM, (c) 64QAM, (d) 256QAM. The considered FFNN scheme operates with 16-bin and 64-bin AHs datasets obtained from mQAM signals with 100 symbols/cluster. The insets show the mean of FFNN-estimated EVMs.

to five does not improve the MAE. With each additional layer, the training time slightly increases due to the complexity of the model.

However, it is still approximately 30 times faster than for the reference CNN structure with similar accuracy. The floating-point operations

(FLOPs) measure the amount of computation for a neural network model [35]. We list Giga (G) FLOPs for each one of those structures in Table 1. Considering the tradeoff between the accuracy and the model complexity, the four-layer FFNN structure with 1000, 500, 500, 100, and 1 neurons is used for further investigations.

First, we illustrate the estimated MAE for different values of N-symbol/cluster and datasets represented as images and vectors (see Fig. 3). Unless 8-bin AHs are used, the MAE below 0.5% can be achieved with only $N = 100$ symbols/cluster. Yet, a substantial difference in estimation accuracy occurs for lower N values. The IQ dataset after CPR provides the most accurate estimation. However, such performance comes at the price of increased complexity. When operating with short sequences before the CPR, the direct input of the vectorized AHs to a

After carefully considering the trade-offs, for further studies, we choose to use AHs with 64 bins.

Finally, we explore how the number of symbols per cluster impacts the MAE for different transmission distances and modulations formats when AHs with 64 bins are used as the input to the FFNN. Figure 5 shows both the estimated MAE and the normalized MAE values for the test datasets. The normalized MAE is calculated by dividing an MAE and its corresponding true EVM value. The normalized MAE allows comparing the estimation accuracy across mQAM modulation formats of a different order. It can be observed that the normalized MAE is below 6.7% (QPSK), 2.4% (16QAM), 2.1% (64QAM), 1.0% (256QAM) using only 100 symbols per cluster. The estimation accuracy of 16QAM, 64QAM, and 256QAM is higher than for the QPSK. This can be explained

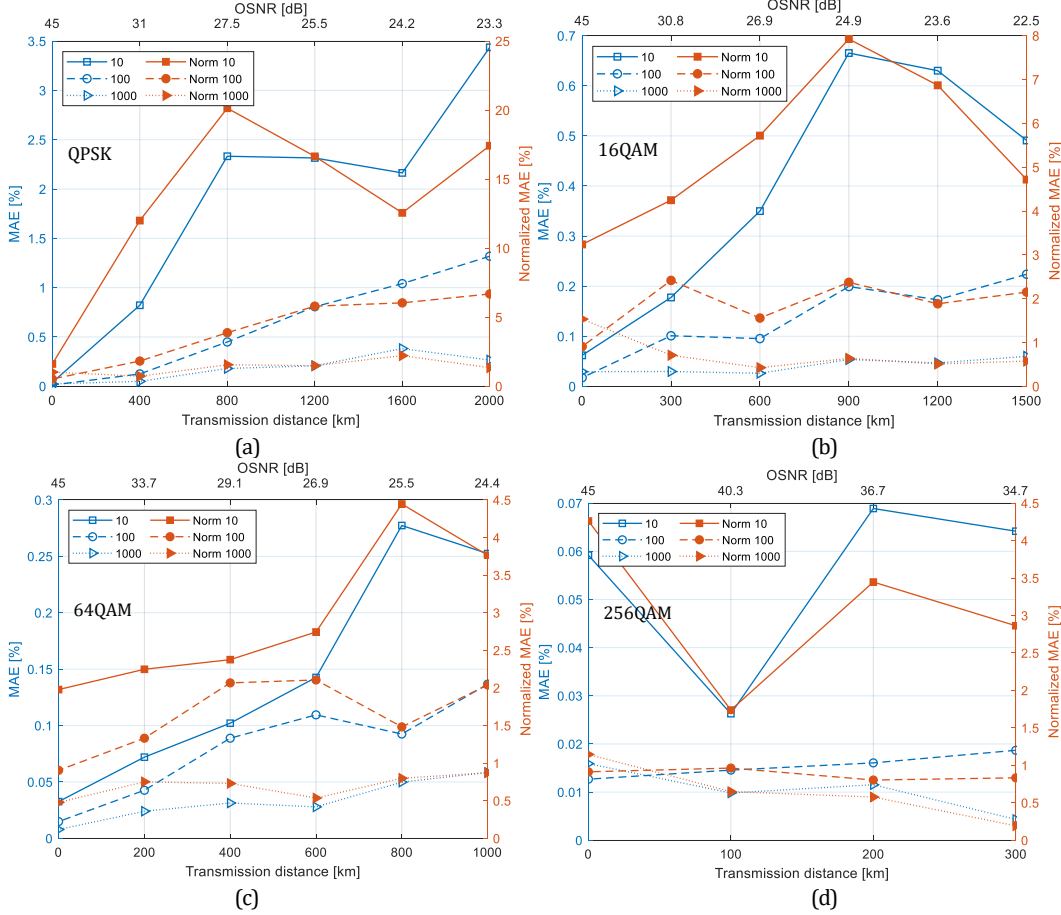


Fig. 5. The MAE performance of the FFNN-based EVM estimator for the test datasets, when it operates with 64-bin AHs containing 10- to 1000-symbol/cluster. (a) QPSK, (b) 16QAM, (c) 64QAM, (d) 256QAM.

neural network ensures a lower MAE compared with the input of images of AHs. At $N = 100$ symbols/cluster, the FFNN with 64-bin AHs can provide comparable performance to the CNN operating with IQs after the CPR.

Next, we evaluate the performance of the FFNN-based scheme when operating with the vectorized AHs having 16 and 64 bins. Figure 4 shows the estimated EVMs for the test datasets and the true EVMs for a specific transmission distance and each considered modulation format. The measured OSNR values are also shown on the upper X-axis of each plot. For 16QAM and 64QAM signals, AHs with 16 and 64 bins allow achieving similar performance. However, a substantial difference is observed for 256QAM signals where a higher bin resolution is beneficial.

by a thresholding effect with respect to the number of clusters. Namely, with the same number of N symbols per cluster, we can get longer sequences captured for a higher-order modulation format and, therefore, more features are fed for EVM estimation. In such a way, more information about the signal quality can be extracted. With the increase of transmission distance, we observe that the estimation results of the 10-symbol/cluster dataset are more fluctuant than they are for the 1000-symbol/cluster dataset. The high estimation variances are caused by the too-short signal sequence.

4. ENERGY CONSUMPTION ANALYSIS

Table 2. Energy consumption for 64-bin AH dataset. Training dataset size: 1650 samples, testing dataset size: 550 samples.

Energy consumption	Training [J]	Testing [J]
CNN	194487.9	290.1
FFNN	2856.7	10.3

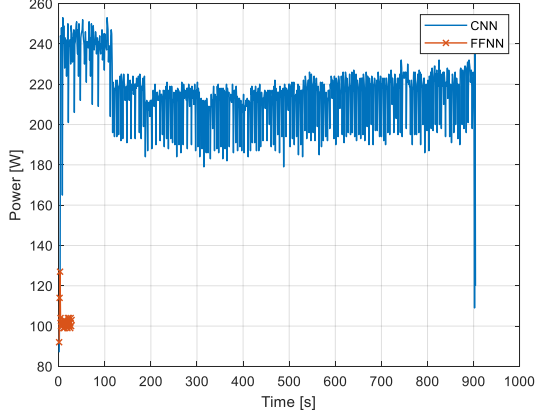


Fig 6. The real-time GPU power for CNN and FFNN based EVM estimator.

The energy consumption of a neural network depends on hardware devices. We run the deep learning-based EVM estimator on NVIDIA's GPUs and therefore we have the possibility to use its System Management Interface (nvidia-smi [36]) tool to monitor the GPU status. We use its option of recording the power usage while running different types/structures of neural networks. Figure 6 shows a real-time GPU power for the CNN and FFNN estimators when the 100-symbols/cluster 64-bin AH dataset is used. It can be observed that the FFNN estimator not only has a shorter operation time but the average GPU power usage is reduced by half. The training and testing energy consumption for the CNN and the FFNN are summarized in Table 2. The low complexity FFNN estimator saves over 95% of the energy compared with the CNN case, making the proposed EVM estimation scheme feasible to low-cost OPM devices distributed in intermediate network nodes.

The overall energy consumption reduction depends on the OPM location. For OPM at the receiver, the monitoring functionality is added to a conventional coherent receiver. In this case, the operational complexity of OPM depends on the ML structure. For an OPM module installed at the intermediate network nodes, full recovery of the signal is not required, and typical metrics used to estimate signal/link quality include the measured signal power and OSNR. Estimating EVM can extend the functionality of the OPM module to directly evaluate the performance of the transmitted signal. Therefore, we aim at minimizing the required DSP process to reduce complexity and save energy, and with the FFNN-based approach, we can completely turn off the CPR block. Further details on the CPR enabling/disabling are left for the implementation.

5. EXPERIMENTAL VALIDATION

We perform the validation of the CNN-based and the FFNN-based schemes for EVM estimation using the experimental setup of 28 Gbaud 64QAM signals (see Fig. 7). First, we use a pseudo-random bit sequence (PRBS) with a word length of $2^{15}-1$ as the input bits. The PRBS bits are

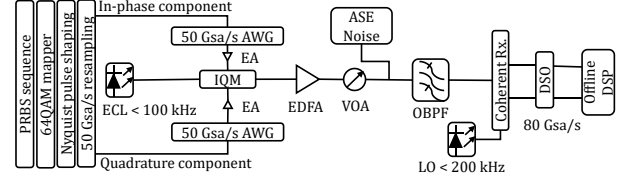


Fig. 7. Experimental setup for 28 Gbaud 64QAM which was used for dataset collection. PRBS, pseudorandom bit sequence; ECL, external cavity laser; EA, electrical amplifier; IQM, in-phase and quadrature modulator; EDFA, erbium-doped fiber amplifier; VOA, variable optical attenuator; ASE, amplified spontaneous emission; OBPF, optical bandpass filter; LO, local oscillator; DSO, digital sampling oscilloscope.

mapped to the symbols to form 64QAM signals. This operation is followed by Nyquist pulse shaping (with a 0.15 roll-off factor) to generate bandwidth-limited signals. Then, resampling of the signal sequences is performed to match the sampling rate of the arbitrary waveform generator (AWG). After that, the electrical signal is amplified in electrical amplifiers (EAs) and sent to an in-phase and quadrature modulator (IQM). We use an external cavity laser (ECL) with 100 kHz linewidth as a continuous wave (CW) light source to obtain the modulated optical signal at the transmitter. Additionally, we use an EDFA to amplify the signal at the transmitter. To adjust the OSNR of the modulated signal, we use a variable optical attenuator (VOA) and two cascaded EDFAs without input signal as an amplified spontaneous emission (ASE) noise source. We obtain the OSNR values from 25 to 44 dB at the resolution of 0.1 nm. This OSNR range allows obtaining the BER values across and below the soft-decision forward error correction

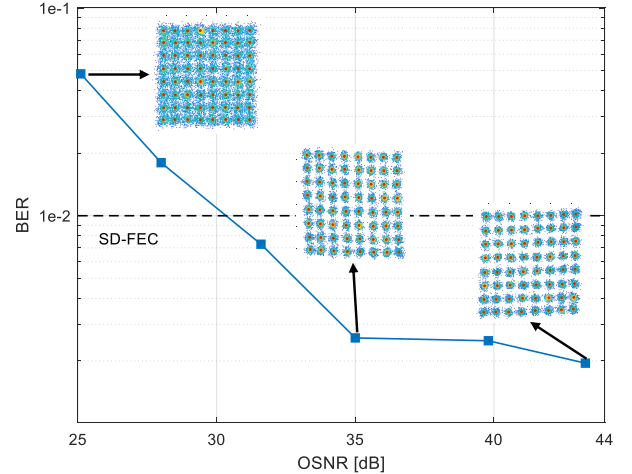


Fig. 8. BER versus OSNR for the considered 28 Gbaud 64QAM signals. The insets are examples of the constellation diagram.

(SD-FEC) limit. After the OSNR loading setup, we place an optical bandpass filter (OBPF) to remove the excess noise from the system not to overload the coherent receiver. In the receiver, we use a local oscillator (LO) laser with a 200 kHz linewidth. Therefore, the combined laser linewidth is around 300 kHz [37]. The coherent beating product between the transmitted signal and the LO laser signal is detected in balanced photodetectors. It is then sampled by an 80 Gsa/s digital sampling oscilloscope (DSO). In Fig. 8, we show the BER vs. OSNR curve for the considered 28 Gbaud 64QAM configuration. The complex signal constellation diagram examples are shown as insets. It can be observed that the signals are limited by the AWGN and the inherent

implementation penalty. We can observe non-perfect square shapes of IQs at high OSNR values.

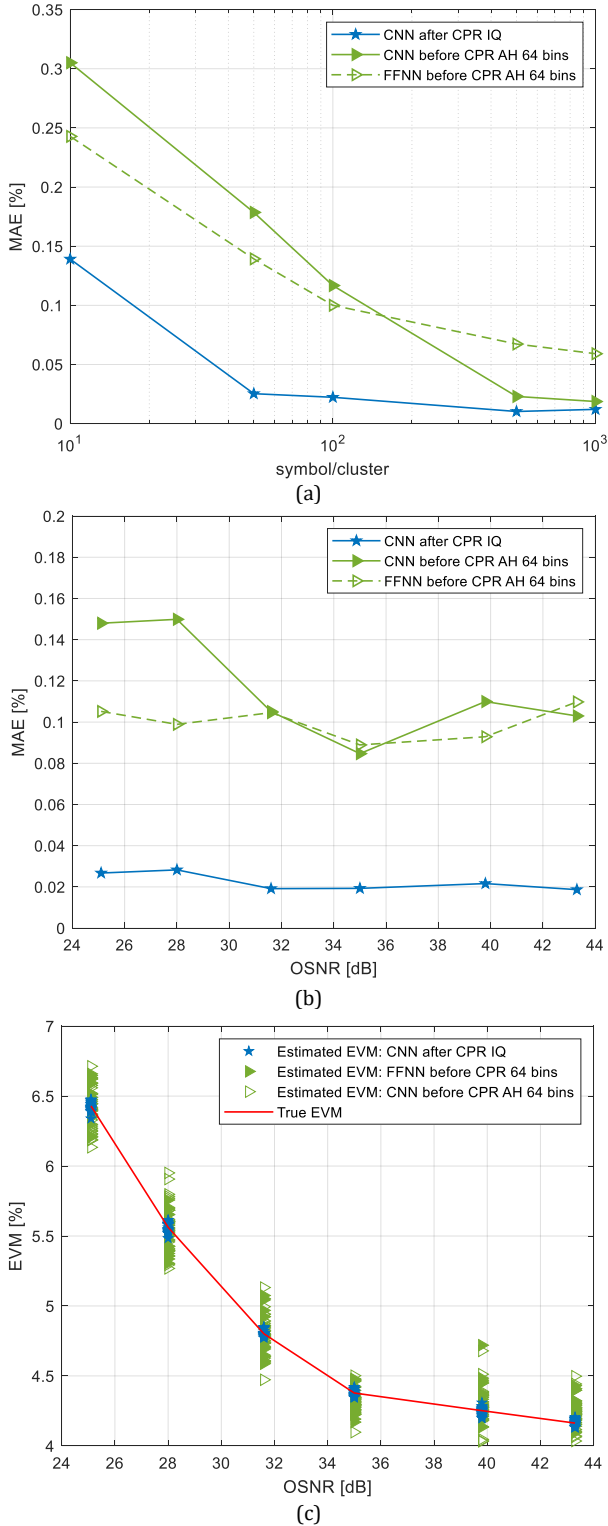


Fig. 9. Experimental validation of the EVM estimator: (a) MAE vs. N symbols per cluster, (b) MAE vs. OSNR, and (c) true EVM vs. estimated EVM for signal sequences of $N = 100$ symbols/cluster.

As before, we collect the images of IQs after the CPR, images of AHs with 64 bins before the CPR as well as their vectorized representations

to form three separate datasets. For this purpose, we use short 64QAM signal sequences with a length of 10 to 1000 symbols per cluster. We split each dataset into training (50%), validation (25%), and testing (25%) subsets. Then we perform the EVM estimation using the CNN- and the FFNN-based schemes. The results are shown in Figure 9. From Fig. 9 (a), one can see that the estimation accuracy increases with the number of symbols per cluster. For a 100-symbol/cluster signal sequence, the FFNN-based scheme achieves 0.1% MAE. Furthermore, the accuracy does not change with the OSNR, as shown in Fig. 9 (b). The CNN-based scheme with IQ images provides an MAE of 0.02%. It is much smaller than the other two cases which come at the cost of more complex signal processing. Finally, Fig. 9 (c) illustrates what is the actual impact of the MAE on the EVM deviation from its true values. For our experimental implementation of the 28 Gbaud 64 QAM signal, the EVM deviation is below 0.5 %. Therefore, the FFNN-based scheme can ensure decent EVM estimation accuracy and provide energy savings thanks to the operation with the vectorized representation of AHs taken before the CPR processing.

6. CONCLUSION

This paper compares and experimentally validates two EVM estimation schemes for signal quality monitoring in coherent communication systems. We exploit CNN- and FFNN-based structures to propose a low-complexity energy-efficient EVM estimation model. Therefore, we also consider the simplified DSP without CPR, which helps to save time and energy required for processing. In the first case, we input images of IQs or AHs obtained from a short signal sequence after or before the CPR, respectively. In the second case, we input the vectorized representation of AHs. Both schemes operate with short signal sequences for time-sensitive EVM monitoring. The performance of the proposed schemes is verified using 32 Gbaud QPSK, 16QAM, 64QAM, and 256QAM signals after 2000km, 1500km, 1000km, and 300km of transmission over the EDFA-amplified fiber-optic link. The estimation accuracy, expressed in terms of MAE and EVM deviation, is studied for different signal lengths, OSNR values, and dataset representations. Compared with the previously proposed centroid-based EVM estimation scheme [9], the proposed FFNN-based scheme particularly relaxes the signal processing requirements by inferring the EVM from the signal representation before the CPR processing. Furthermore, we calculated that the FFNN-based scheme consumes by 95% less energy as compared with the CNN-based solution. Besides, experimental validation is performed using the 28 Gbaud 64QAM configuration. With both schemes, we achieve the MAE below 0.15% for a wide range of OSNR values when operating with the signal sequence of 100 symbols per cluster. Therefore, these results indicate that the proposed schemes can be used to perform time-sensitive and accurate EVM estimation for mQAM signal quality monitoring in coherent communication systems.

Acknowledgment.

This work was supported by the Vetenskapsrådet projects PHASE (2016-04510) and 2019-05197, and RISE SK funded project Optical Neural Networks (P109599).

References

1. K. Roberts, Q. Zhuge, I. Monga, S. Gareau and C. Laperle, "Beyond 100 Gb/s: capacity, flexibility, and network optimization," in *IEEE/OSA J. Opt. Commun. and Netw.* **9**, C12-C23 (2017).
2. E. Maniloff, S. Gareau and M. Moyer, "400G and Beyond: Coherent Evolution to High-Capacity Inter Data Center Links," 2019 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2019, pp. 1-3.

3. D. C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Ostar, M. Preiss, and A. E. Willner, "Optical performance monitoring," *J. of Lightwave Technol.* **22**, 294-304 (2004).
4. Z. Pan, C. Yu, and A. E. Willner, "Optical performance monitoring for the next generation optical communication networks," *Optical fiber technology* **16**, 20-45 (2010).
5. F. N. Khan, Z. Dong, C. Lu, and A. P. T. Lau, "Optical performance monitoring for fiber-optic communication networks," in *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks*, John Wiley & Sons, 2016, Chap. 14.
6. R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Hübner, C. Koos, J. Becker, W. Freude, and J. Leuthold, "Error vector magnitude as a performance measure for advanced modulation," *IEEE Photon. Technol. Lett.* **24**, 61-63 (2012).
7. I. Fatadin, "Estimation of BER from Error Vector Magnitude for Optical Coherent Systems," *Photonics* **3**, p. 21 (2016).
8. R. Schmogrow, D. Hillerkuss, S. Wolf, B. Bäuerle, M. Winter, P. Kleinow, B. Nebendahl, T. Dippon, P. C. Schindler, C. Koos, W. Freude, and J. Leuthold, "512QAM Nyquist sinc-pulse transmission at 54 Gbit/s in an optical bandwidth of 3 GHz," *Opt. Express* **20**, 6439-6447 (2012).
9. Y. Fan, A. Udalcovs, X. Pang, C. Natalino, M. Furdek, S. Popov, and O. Ozolins, "Fast signal quality monitoring for coherent communications enabled by CNN-based EVM estimation," *J. Opt. Commun. Netw.* **13**, B12-B20 (2021).
10. Y. Fan, A. Udalcovs, X. Pang, C. Natalino, R. Schatz, M. Furdek, S. Popov, and O. Ozolins, "Deep Learning Assisted Pre-Carrier Phase Recovery EVM Estimation for Coherent Transmission Systems," in 2021 Conference on Lasers and Electro-Optics (CLEO), OSA Technical Digest (online) (Optical Society of America, 2021), accepted STh1F.2.
11. X. Chen, B. Li, R. Proietti, C. Liu, Z. Zhu, and S. J. B. Yoo, "Demonstration of distributed collaborative learning with end-to-end QoT estimation in multi-domain elastic optical networks," *Opt. Express* **27**, 35700-35709 (2019).
12. A. P. Vela, M. Ruiz, F. Fresi, N. Sambo, F. Cugini, G. Meloni, L. Poti, L. Velasco, and P. Castoldi, "BER Degradation Detection and Failure Identification in Elastic Optical Networks," *J. Lightwave Technol.* **35**, 4595-4604 (2017).
13. C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," *J. Opt. Commun. Netw.* **10**, A286-A297 (2018).
14. C. Wang, S. Fu, Z. Xiao, M. Tang, and D. Liu, "Long Short-Term Memory Neural Network (LSTM-NN) Enabled Accurate Optical Signal-to-Noise Ratio (OSNR) Monitoring," *J. Lightwave Technol.* **37**, 4140-4146 (2019).
15. T. Tanimura, T. Hoshida, T. Kato, S. Watanabe, and H. Morikawa, "Convolutional Neural Network-Based Optical Performance Monitoring for Optical Transport Networks," *J. Opt. Commun. Netw.* **11**, A52-A59 (2019).
16. X. Fan et al., "Feature Fusion-Based Multi-Task ConvNet for Simultaneous Optical Performance Monitoring and Bit-Rate/Modulation Format Identification," *IEEE Access* **7**, pp. 126709-126719 (2019).
17. Z. Wang, A. Yang, P. Guo, and P. He, "OSNR and nonlinear noise power estimation for optical fiber communication systems using LSTM based deep learning technique," *Opt. Express* **26**, 21346-21357 (2018).
18. D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Express* **25**, 17150-17166 (2017).
19. 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. Express* **25**, 17767-17776 (2017).
20. T. S. R. Shen, K. Meng, A. P. T. Lau and Z. Y. Dong, "Optical Performance Monitoring Using Artificial Neural Network Trained with Asynchronous Amplitude Histograms," *IEEE Photon. Technol. Lett.* **22**, 1665-1667 (2010).
21. C. Natalino, A. Udalcovs, L. Wosinska, O. Ozolins, and M. Furdek, "One-shot learning for modulation format identification in evolving optical networks," in OSA APC (IPR, Networks, NOMA, SPPCom, PVLED), OSA Technical Digest (Optical Society of America, 2019), paper JW4A.2.
22. F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications," *J. Lightwave Technol.* **37**, 493-516 (2019).
23. F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "Machine learning methods for optical communication systems and networks," in *Optical Fiber Telecommunications*, 7th ed. (Academic, 2019), Chap. 21.
24. T. Kupfer, A. Bisplinghof, T. Duthel, C. Fludger, and S. Langenbach, "Optimizing Power Consumption of a Coherent DSP for Metro and Data Center Interconnects," in Optical Fiber Communication Conference, OSA Technical Digest (online) (Optical Society of America, 2017), paper Th3G.2.
25. VPIphotonics GmbH, "VPItransmissionMaker11," 2021, <https://www.vpiphotonics.com/>.
26. Q. Zhang, Y. Yang, C. Guo, X. Zhou, Y. Yao, A. P. T. Lau, and C. Lu, "Accurate BER estimation scheme based on K-means clustering assisted Gaussian approach for arbitrary modulation format," *J. Lightwave Technol.* **38**, 2152-2157 (2020).
27. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. (Springer Series in Statistics Springer New York Inc., New York, 2009), Chap. 11.
28. D. Svozil, V. Kvasnicka and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemometrics and intelligent laboratory systems* **39**, 43-62 (1997).
29. X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural network," *J. Mach. Learn. Res.* **15**, 315-323 (2011).
30. D. E. Rumelhart and J. L. McClelland, "Learning Internal Representations by Error Propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*, MIT Press, 1987, pp.318-362.
31. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. (Springer Series in Statistics Springer New York Inc., New York, 2009), Chap. 2.
32. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in International Conference for Learning Representations (ICLR), California, USA (2015).
33. F. Chollet, "Keras," 2021, <https://keras.io>.
34. M. Abadi, "TensorFlow: large-scale machine learning on heterogeneous distributed systems," in Proc. Conf. Lang. Resour. Eval., March 2016, pp. 3243-3249.
35. P. Molchanov, S. Tyree, T. Karras, T. Aila, and J. Kautz, "Pruning convolutional neural networks for resource efficient inference," in Proceedings of the International Conference on Learning Representations (ICLR), 2017.
36. NVIDIA system management interference (nvidia-smi) tool, 2021, <https://developer.nvidia.com/nvidia-system-management-interface>.
37. J. R. Navarro, A. Udalcovs, X. Pang, O. Ozolins, A. Kakkar, R. Schatz, F. Nordwall, S. Popov, and G. Jacobsen, "High Phase Noise Tolerant Circular-64QAM with Efficient Phase Recovery for Coherent Optical Systems," in Asia Communications and Photonics Conference, OSA Technical Digest (online) (Optical Society of America, 2017), paper Su3B.7.