# **MENTOR** Machine LEarning in optical NeTwORks

A European Industrial Doctorate [EC GA 956713]

## DELIVERABLE 1.2 (D2)

# The augmentation strategy for the MB optical line elements

---

## Document Details

| | |
|---|---|
| Title | The augmentation strategy for the MB optical line elements |
| Deliverable number | Deliverable 1.2 |
| Deliverable Type | Public |
| Work Package | WP1 |
| Deliverable due date | 1 September 2023 (Postponed from 31 Mar 2023) |
| Actual date of submission | 7 August 2023 |
| Lead beneficiary | Aston University |
| Main Author | Sasipim Srivallapanondh |

## Project Details

| | |
|---|---|
| Grant Agreement | 956713 |
| Project Acronym | MENTOR |
| Project Title | Machine LEarning in optical NeTwORks |
| Call Identifier | H2020-MSCA-ITN-2020 |
| Project Website | https://mentor.astonphotonics.uk |
| Start of the Project | 1 January 2021 |
| Project Duration | 48 months |

## Dissemination level

| | | |
|---|---|---|
| PU | Public | X |
| CO | Confidential, only for members of the consortium (including Commission Services | |

## EC Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 956713

# Table of Contents

# List of Figures

# Beneficiaries

Aston Institute Of Photonic Technologies (Lead Beneficiary at Aston University)
Infinera Germany (Non-recruiting Beneficiary)
Infinera Portugal (Non-recruiting Beneficiary)
Sant'Anna School of Advanced Studies
Technical University of Denmark
Polytechnic University of Catalonia

# Industrial Partners

Telecom Italia Mobile
Orange Telecom Company

# List of Acronyms

| | |
|---|---|
| AiPT | Aston Institute Of Photonic Technologies |
| AST | Aston University |
| CA | Consortium Agreement |
| DTU | Technical University of Denmark |
| EC | European Commission |
| EID | European Industrial Doctorate |
| ESR | Early Stage Researcher |
| GA | Grant Agreement |
| INF-G | INFINERA Germany (Non-recruiting Beneficiary) |
| INF-P | INFINERA Portugal (Non-recruiting Beneficiary) |
| MENTOR | Machine LEarning in optical NeTwORks |
| ORANGE | Orange Telecom Company |
| SB | Supervisory Board |
| SSSA | Sant'Anna School of Advanced Studies |
| TIM | Telecom Italia Mobile |
| UPC | Polytechnic University of Catalonia |

# Executive Summary

**Overview**

This deliverable recapitulates the data augmentation techniques and the concept of multi-band (MB) optical networks. The general description of data augmentation and its features are presented. The data augmentation technique is not only promising for generating synthetic data but also helps reduce the training complexity of neural networks (NN). In particular, the investigation focuses on the data augmentation used in NN-based optical equalisers. It has been proven that with data augmentation, the dataset size required is reduced. The data used in MB networks can be limited because it is beyond the most common band (C-band). Therefore, data augmentation becomes useful to generate synthetic data.

# 1. NN-Based Equalisers and its Challenges

The demand for high-speed data transmission keeps increasing due to upcoming technologies (6G [1], etc.). Coherent optical systems have emerged as a key solution to meet this demand. Nonetheless, the presence of linear and especially nonlinear distortions in fibre-optic systems limits the achievable information rates [2, 3, 4]. Various digital signal processing (DSP) techniques have been proposed for nonlinear effects mitigation in long-haul systems [3]. Neural networks (NNs) have recently emerged as an effective alternative for channel equalisation: the NNs have demonstrated excellent capability to approximate the inverse of the optical channel transfer function, potentially outperforming conventional DSP approaches [5, 6, 7]. NNs are intrinsically nonlinear, thus they work well with the type of effects that occur in optical networks. Figure 1 is an example of the NN-based equaliser's diagram. However, major challenges still arise, for example, high computational complexity and low generalisability. To reduce the computational complexity, there are both direct and indirect ways and both in the training and the inference phases. For the training phase, reducing complexity is crucial for the efficient and practical deployment of hardware with limited resources. We can obtain the lower complexity by various methods, namely, transfer learning or approaches to improve generalisation, such as data augmentation, domain randomisation, and semi-supervised learning. For the inference phase, different techniques to reduce the computational complexity are: pruning, quantisation, knowledge distillation and weight clustering. In this report, we focus on data augmentation.

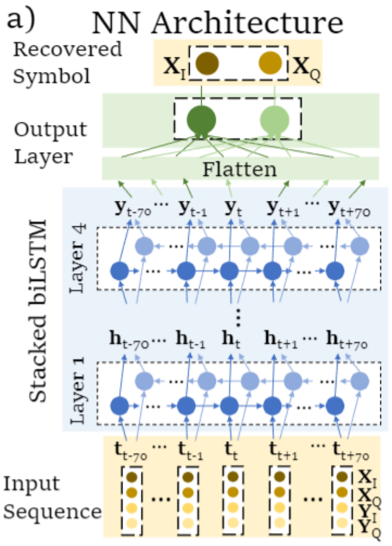

*Figure 1: An example of the NN-based equaliser.*

# 2. Data Augmentation

Data augmentation allows datasets to be more diverse and representative by artificially generating additional data points from existing data *[8]*. This method increases the size of a training dataset by applying various transformations to the existing data. With more diverse training data points, the model can the better performance and generalisation, resulting in a less-overfitting model. To provide an obvious example of data augmentation, we can consider an image as our data. Data augmentation techniques of the images can be rotation, cropping, flipping, colour space and noise injection *[9]*. Figure 2 illustrates the images with the data augmentation techniques. For the optical fibre network, data augmentation has also been applied to different applications, such as failure management *[10]* and nonlinear mitigation *[11]*.

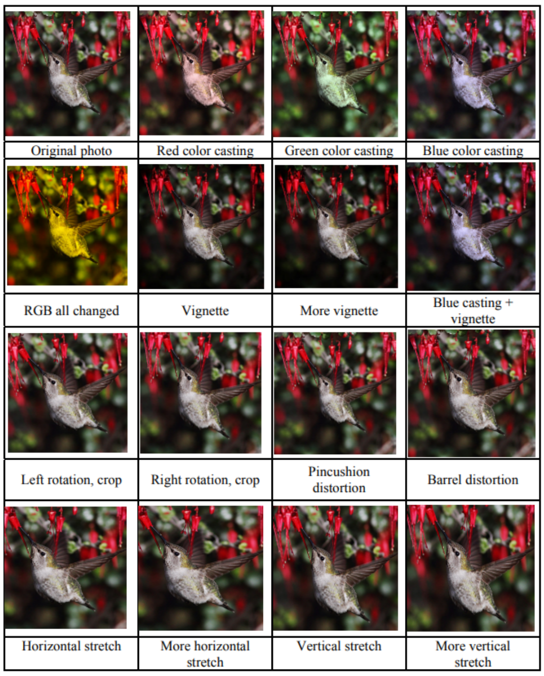

*Figure 2: Examples of colour augmentations tested by [12].*

## Features of Data Augmentation

1. To increase the training data: Many applications do not have access to big data, and with the supervised training, it is data-driven. With data augmentation, the number of training samples can be expanded significantly. It is beneficial especially when the original dataset is limited or imbalanced. More training data assists the model in capturing a broader range of patterns.
2. To improve the robustness: Data augmentation introduces variations to the training data. The different variations in data help the model to be more robust to changes in the unseen data, improving the generalisability.
3. To reduce the overfitting: With data augmentation, the model learns to focus on the important features and becomes less sensitive to irrelevant variations in the data because by adding noises to the data, the model learns more diversity in the training data.
4. To reduce the training complexity: Data augmentation can indirectly reduce the training complexity. Ref. [10] shows a modified training dataset can reduce the training time of NN. The model can converge faster when the training data is effective. The augmented data provides a more comprehensive representation of the real-world variations, allowing the model to generalise better with a potentially simpler model structure.
5. To save computational resources in the inference phase: Ref. [10] demonstrates that a good quality training dataset obtained using data augmentation can save computational resources.

## Data Augmentation in NN-based Equalisers

Data augmentation used in optical NN-based equalisers is a technique to improve equalisation performance and decrease the training complexity of supervised learning in nonlinearity mitigation. In the supervised learning tasks, normally a large training dataset is required. The model will also need to be re-trained when the channel conditions change. However, big data collection can be challenging. The efficient use of a limited dataset is more desirable for practical implementation. It has been demonstrated that data augmentation reduces the size of the dataset up to 6 times while maintaining the optical performance *[11]*. Data augmentation enhances the diversity of the dataset by generating new artificial input samples from the original ones. The model for nonlinear mitigation in optical coherent transmission systems with data augmentation has proven to require fewer parameters and, consequently, converges faster, as shown in *[11]* for the NN equalisers. Note that data augmentation is also applied in other sections of optical communications, for example, predicting failures *[13]*, failure management *[10]* and traffic peculiarities *[14]*.

In nonlinear mitigation, the NN aims to predict the transmitted signal $u(0,t)$, from the received one $u(z,t)$. The dataset then follows the Manakov equations which describe the averaged evolution of a dual-polarised (DP) optical signal along an optical fiber link:

$$\frac{\partial u_{h/v}}{\partial z} = \frac{-\alpha(z)}{2} u_{h/v} - i\frac{\beta_2(z)}{2}\frac{\partial^2 u_{h/v}}{\partial t^2} + i\frac{8\gamma(z)}{9}(|u_h|^2 + |u_v|^2)u_{h/v} + \xi(z,t)$$

Where $u_h(z,t)$ and $u_v(z,t)$ are the horizontal (h) and vertical (v) polarisations of the optical signal waveform $u(z,t)$, $\alpha(z)$ is the attenuation, $\beta_2(z)$ is the group velocity dispersion (GVD) coefficient, $\gamma(z)$) is the effective nonlinear coefficient and $\xi(z,t)$ is the amplified spontaneous emission (ASE) noise injected by optical amplifiers.

For the data augmentation in this case from Ref. [11], the signal is transformed by three methods:

1. the discrete phase shift: $(\Delta\varphi_{disc}) - \{\bar{u}_{h/v}(0,t)exp^{i\varphi}, \bar{u}_{h/v}(z,t)exp^{i\varphi}\}, \forall\varphi \in \left\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\right\}$
2. the time-inversion: $(t_{inv}) - \{u_{h/v}(0,-t), u_{h/v}(z,-t)\}$
3. the polarisation swapping: $(H/V_{swap}) - \{u_{u\leftrightarrow v}(0,t), u_{h\leftrightarrow v}(z,t)\}$.

Ref. [11] proposed to do the transformation of the dataset at every training epoch by replacing a randomly chosen part of the original dataset with the data points generated by the data augmentation methods mentioned. A randomly chosen discrete phase shift $\Delta\varphi_{disc}$ is applied to every dataset object, while the time-inversion $t_{inv}$ and polarisation swapping $H/V_{swap}$ are applied separately to randomly chosen halves of the dataset objects.

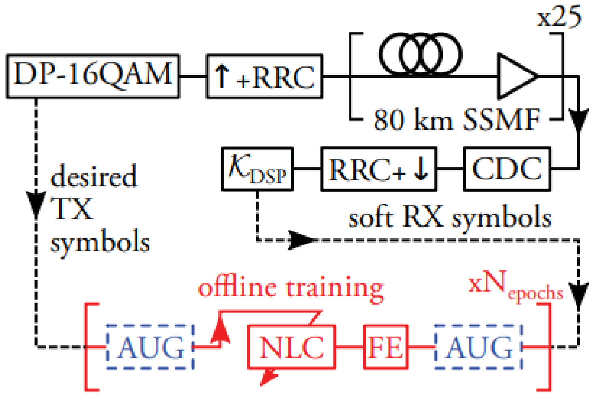

*Figure 3: Data augmentation in numerical setup [11].*

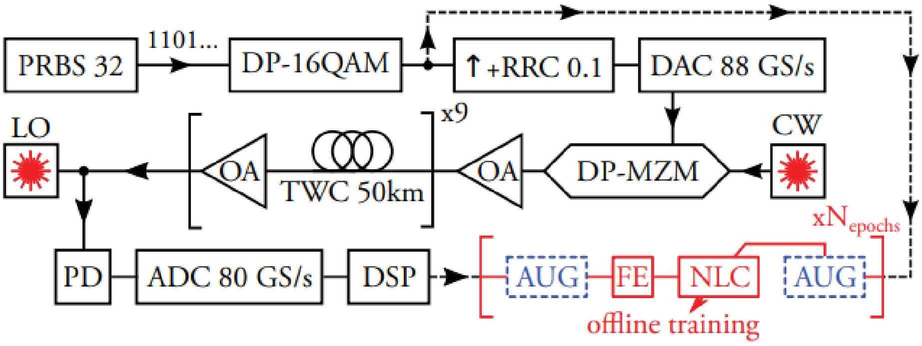

*Figure 4: Data augmentation in experimental setup [11].*

Figure 3 and Figure 4 show the integration of the data augmentation scheme in the numerical and experimental setup, respectively. To augment the received signal features, like nonlinear perturbation triplets (NPTs), one can, the transformations (AUG) in Figure 3 can be applied to the received signal $u(z,t)$ and then perform the feature extraction (FE). It is worth noting that this data augmentation is used after conventional digital signal processing (DSP).

Their results in Figure 5 (numerical) and Figure 6 (experimental) show the BER for different training data size ($N_{tr}$) without data augmentation (pure data) and with data augmentation (joint aug). The dashed line illustrated the BER before nonlinear compensation (NLC). Note that in the experimental study, the time-inversion was excluded.

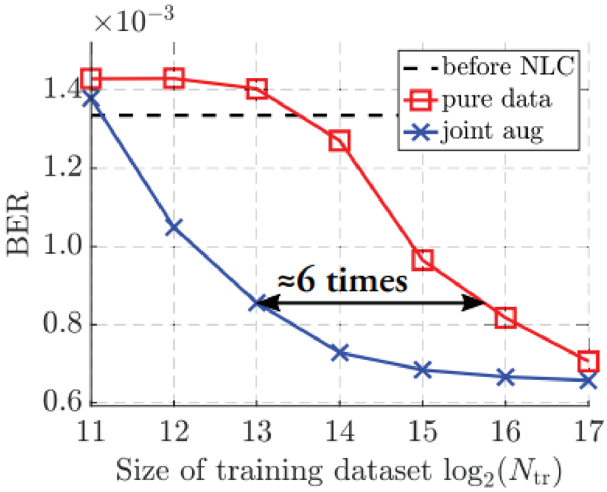

*Figure 5: Numerical study of the dependence of BER (after SL-NLC) on the size of dataset [11].*

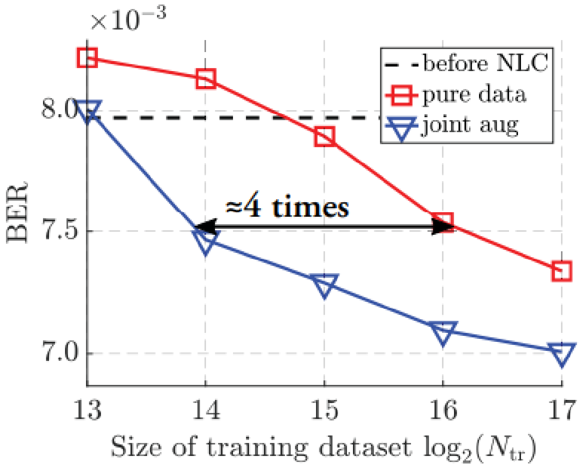

*Figure 6: Experimental study of the dependence of BER after SL-NLC on the size of dataset [11].*

For the numerical case, the joint augmentation showed the same BER achieved when the dataset is 6 times smaller than the pure dataset. In the experiment, the reduction of the dataset size by 4 times is obtained by the joint augmentation and it still maintains the same BER after the DNN [11].

# 3. Multi-Band Optical Networks

Due to the increasing IP traffic demand resulting from 5G, high-capacity access traffic and cloud services, the backbone optical network infrastructure can be loaded. Therefore, the per-fibre throughput enabled by the coherent transmission technologies is no longer sufficient to support the envisioned IP traffic explosion [15]. To upgrade the infrastructures to deal with the insufficient per-fibre throughput without deploying new fibres, multi-band (MB) transmission can be a promising solution. MB maximises the per-fibre transmission by exploiting spectral bands beyond the C-band which are O, E, S, and L bands. MB aims to transmit the signal over the entire low-loss optical spectrum of single-mode fibres (SMF), which is in the range from 1260 nm to 1625 nm. With different bands, the nonlinearity and the fibre loss vary.

MB networks utilise multiple optical bands to transmit signals through wavelength division multiplexing in SMF. This technology enables the networks to achieve an optical bandwidth up to ten times greater than conventional systems, which are limited to using only the C-band. The primary objective of MB is to maximise the return on investment for existing optical infrastructure. Already, the extension to the L-band is commercially available [16]. Despite the degradation of the generalised optical signal-to-noise ratio (GSNR) caused by using more bands in the C-band, MB-based systems can handle larger data traffic [17]. However, to expand to other bands, factors like fibre loss, dispersion, and nonlinearity of each band need to be carefully considered. In Figure 9, the dispersion and fibre loss characteristics of each band are depicted, revealing that the C and L bands have the lowest fibre loss, while the E band exhibits the least dispersion.

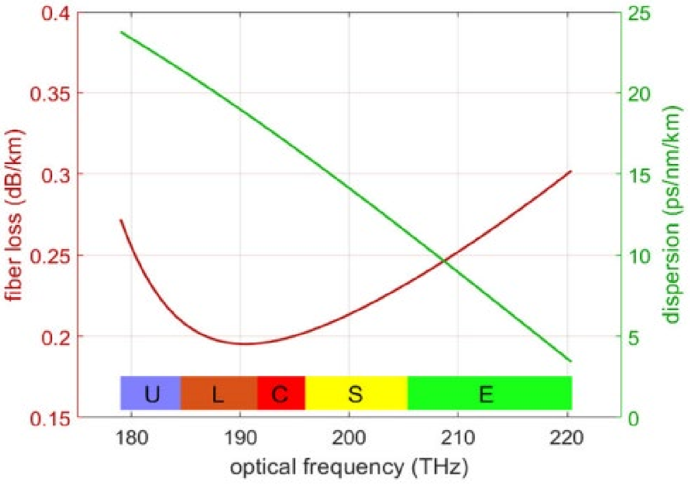

*Figure 7 : Loss (red line) and dispersion (green line) characteristics in each optical band [16].*

When considering throughput, the addition of more bands has the potential to increase overall data transfer rates. However, it is important to note that simply widening the band may not always guarantee higher capacity due to limitations imposed by nonlinearity and fibre loss. *Figure 6*, as referenced in *[16]*, illustrates the relationship between total throughput, channel number, and the contribution of the next-channel information rate. The combined C+L band offers higher throughput. Incorporating the S-band appears to increase throughput efficiently. Nonetheless, the attractiveness of the E-band diminishes after accounting for its range, as the added channels do not significantly enhance throughput. Notably, the O-band did not demonstrate a noticeable improvement in throughput, aligning with similar trends found in other studies *[18] [19]*.

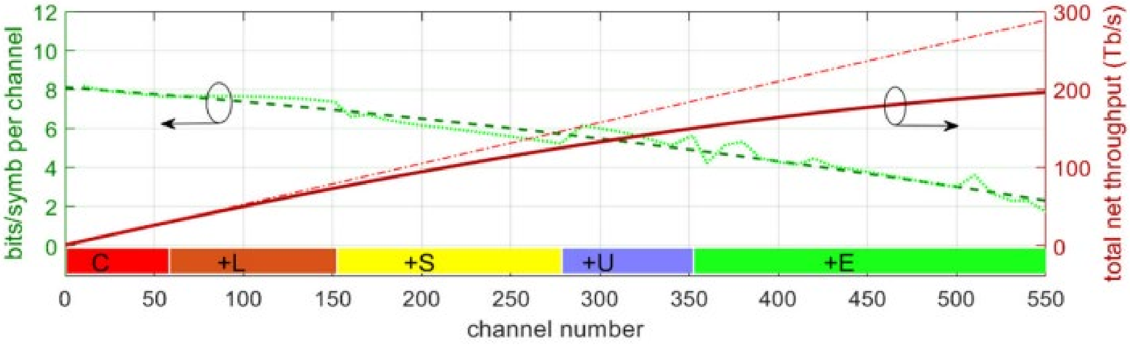

*Figure 8: Red solid line: total system throughput vs. number of WDM channels at the end of a 10x100km SMF link. Channels are progressively added in the bands shown by the band indicator at the bottom of the figure. Red dashed line: linear extrapolation with same slope as for the C band. Green dots: bits/symbol carried by the next added channels; dashes: smoothed version [16].*

Ref. [20] shows the simulation of MB. The entire physical configuration of the multiband setup is depicted in Figure 9. To generate the WDM (wavelength division multiplexing) comb, a 32-GBaud 16QAM transmitter is employed, encompassing the L-, C-, and S- frequency bands. Each pulse within the system has a bandwidth of 60 GHz, and a raised cosine filter with a roll-off factor of 0.06 is utilised. The propagation of the optical signal as it travels through the optical fibre is controlled by the Generalized Non-Linear Schrödinger Equation (GNLSE).

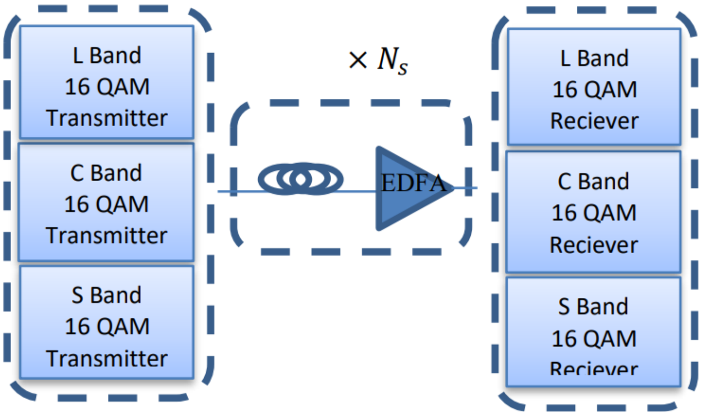

*Figure 9: Physical layout of the multiband setup*

## 4. Data Augmentation for Multi Band Optical Networks

The data collection in MB optical networks is computationally expensive [20] in the simulation and in the experimental, it is still not as available as in the original C-band. The modelling of MB consists of both linear and non-linear effects. The researchers have proposed a variety of MB simulators using different models to simulate the SRS effects and other nonlinear impairments, for example, Gaussian Noise model and Split step methods. Split-step Fourier Method (SSFM) is one of the most popular models to solve the nonlinear Schrodinger equation in optical networks. When there is an increment in bandwidth, the SSFM method can be more challenging due to smaller step size requirements and increased computation complexity [20]. However, to work with machine learning or NN, a large enough dataset is crucial for a data-driven model. With the computational complexity and the limitation in the data generation/collection in the MB system, data augmentation can be a promising solution to increase the dataset size in MB systems.

# CONCLUSION

Data augmentation has shown to be one promising way in generating synthetic data and in reducing the training complexity of the NN-based equalisers. It is very powerful in decreasing the dataset size for the NN to train on and still results in the same level of performance. The benefits of data augmentation have been presented. With the MB structure, more nonlinearity is introduced. Adding more bands can lead to more throughput but it is not necessary that the wider band provides more throughput due to the nonlinearity and fibre attenuation. The training of the data-driven NN-based equalisers can be more challenging. Also, data collection in MB optical networks is computationally expensive. Therefore, data augmentation can be a promising solution to increase the dataset size in MB systems and reduce the training complexity.

# References

[1] H. Holma, H. Viswanathan and P. Mogenses, “Extreme massive mimo for macro cell capacity boost in 5g-advanced and 6g,” *White Paper,* 2021.

[2] R.-J. Essiambre and R. W. Tkach, “Capacity trends and limits of optical communication networks,” *Proceedings of the IEEE,* vol. 100, no. 5, pp. 1035-1055, 2012.

[3] Cartledge, John C., Fernando P. Guiomar, Frank R. Kschischang, Gabriele Liga, and Metodi P. Yankov, “Digital signal processing for fiber nonlinearities,” *Optics express,* vol. 25, pp. 1916-1936, 2017.

[4] P. Bayvel, R. Maher, T. Xu, G. Liga, N. A. Shevchenko, D. Lavery, A. Alvarado and R. I. Killey, “Maximizing the optical network capacity,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* vol. 374, no. 2062, 2016.

[5] F. N. Khan, Q. Fan, C. Lu and A. P. T. Lau, “An optical communication's perspective on machine learning and its applications,” *Journal of Lightwave Technology,* vol. 37, no. 2, pp. 493-516, 2019.

[6] C. Hager and H. D. Pfister, “Nonlinear interference mitigation via deep neural networks,” in *Optical fiber communication conference*, 2018.

[7] Freire, Pedro J., Yevhenii Osadchuk, Bernhard Spinnler, Antonio Napoli, Wolfgang Schairer, Nelson Costa, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, “Performance versus complexity study of neural network equalizers in coherent optical systems,” *Journal of Lightwave Technology,* vol. 39, pp. 6085-6096, 2021.

[8] K. Alomar, H. I. Aysel and X. Cai, “Data Augmentation in Classification and Segmentation: A Survey and New Strategies,” *Journal of Imaging,* vol. 9, no. 2, p. 46, 2023.

[9] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data,* vol. 6, no. 60, 2019.

[10] L. Z. Khan, J. Pedro, N. Costa, L. De Marinis, A. Napoli and N. Sambo, “Data augmentation to improve performance of neural networks for failure management in optical networks,” *Journal of Optical Communications and Networking,* vol. 15, no. 1, pp. 57-67, 2023.

[11] V. Neskorniuk, P. J. Freire, A. Napoli, B. Spinnler, W. Schairer, J. E. Prilepsky, N. Costa and S. K. Turitsyn, “Simplifying the Supervised Learning of Kerr Nonlinearity Compensation Algorithms by Data Augmentation,” in *2020 European Conference on Optical Communications (ECOC)*, 2020.

[12] R. Wu, S. Yang, Y. Yi, Q. Dang and G. Sun, “Deep image: scaling up image recognition,” 2015.

[13] L. Cui, Y. Zhao, B. Yan, D. Liu and J. Zhang, “Deeplearning-based failure prediction with data augmentation in optical transport networks,” in *17th International Conference on Optical Communications and Networks (ICOCN2018)*, 2019.

[14] J. Li, D. Wang, S. Li, M. Zhang, C. Song and X. Chen, “Deep learning based adaptive sequential data augmentation technique for the optical network traffic synthesis,” *Optics express,* vol. 27, no. 13, pp. 18831-18847, 2019.

[15] Alessio Ferrari, Antonio Napoli, Johannes K Fischer, Nelson Costa, Andrea D’Amico, João Pedro, Wladek Forysiak, Erwan Pincemin, Andrew Lord, Alexandros Stavdas, Juan Pedro F-P Gimenez, Gunther Roelkens, Nicola Calabretta, Silvio Abrate, Bernd Sommerkorn-K, “Assessment on the achievable throughput of multi-band ITU-T G. 652. D fiber transmission systems,” *Journal of Lightwave Technology,* pp. 4279-4291, 2022.

[16] Poggiolini, Pierluigi, and Mahdi Ranjbar-Zefreh, “Closed Form Expressions of the Nonlinear Interference for UWB Systems,” *ECOC,* 2022.

[17] Sambo, Nicola, Alessio Ferrari, Antonio Napoli, Nelson Costa, João Pedro, Bernd Sommerkorn-Krombholz, Piero Castoldi, and Vittorio Curri, “Provisioning in multi-band optical networks,” *Journal of Lightwave Technology,* vol. 38, pp. 2598-2605, 2020.

[18] Semrau, Daniel, “Modeling of Fiber Nonlinearity in Wideband Transmission,” in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2022.

[19] D’Amico, Andrea, Bruno Correia, Elliot London, Emanuele Virgillito, Giacomo Borraccini, Antonio Napoli, and Vittorio Curr, “Scalable and disaggregated GGN approximation applied to a C+ L+ S optical network,” *Journal of Lightwave Technology,* vol. 40, pp. 3499-3511, 2022.

[20] P. Khare, N. Costa, J. Pedro, A. Napoli, F. Arpanaei, J. Comellas, M. Ruiz and L. Velasco, “SSMS: A Split Step MultiBand Simulation Software,” in *International Conference on Transparent Optical Networks*, Bucharest, 2023.

[21] Okamoto, Seiji, Kengo Horikoshi, Fukutaro Hamaoka, Kyo Minoguchi, and Akira Hirano, “5-band (O, E, S, C, and L) WDM transmission with wavelength adaptive modulation format allocation,” in *European Conference on Optical Communication*, 2016.

[22] Molchanov, P., S. Tyree, T. Karras, T. Aila, and J. Kautz, “Pruning convolutional neural networks for resource efficient inference,” in *International Conference on Learning Representations*, 2019.

[23] Freire, Pedro J., Yevhenii Osadchuk, Bernhard Spinnler, Antonio Napoli, Wolfgang Schairer, Nelson Costa, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, “Performance versus complexity study of neural network equalizers in coherent optical systems,” *Journal of Lightwave Technology,* vol. 39, pp. 6085-6096, 2021.

[24] Zhu, Michael, and Suyog Gupta, “To prune, or not to prune: exploring the efficacy of pruning for model compression,” *arXiv preprint arXiv:1710.01878,* 2017.

[25] Blalock, Davis, Jose Javier Gonzalez Ortiz, Jonathan Frankle, and John Guttag, “What is the state of neural network pruning?,” *Proceedings of machine learning and systems 2,* pp. 129-146, 2020.

[26] Frankle, Jonathan, and Michael Carbin, “The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks,” *International Conference on Learning Representations,* 2018.

[27] Koike-Akino, Toshiaki, Ye Wang, Keisuke Kojima, Kieran Parsons, and Tsuyoshi Yoshida, “Zero-multiplier sparse dnn equalization for fiber-optic qam systems with probabilistic amplitude shaping,” *European Conference on Optical Communication (ECOC),* 2021.

[28] Napoli, Antonio and Costa, Nelson and Fischer, Johannes K and Pedro, Joao and Abrate, Silvio and Calabretta, Nicola and Forysiak, Wladek and Pincemin, Erwan and Gimenez, Juan PF-P and Matrakidis, Chris and others, “Towards multiband optical systems,” *Photonic Networks and Devices,* 2018.

[29] Liang, Tailin, John Glossner, Lei Wang, Shaobo Shi, and Xiaotong Zhang, “Pruning and quantization for deep neural network acceleration: A survey,” *Neurocomputing,* pp. 370-403, 2021.

[30] Chuang, Chun-Yen, Wei-Fan Chang, Chia-Chien Wei, Ching-Ju Ho, Cheng-Yu Huang, Jin-Wei Shi, Lindor Henrickson, Young-Kai Chen, and Jyehong Chen, “Sparse volterra nonlinear equalizer by employing pruning algorithm for high-speed PAM-4 850-nm VCSEL optical interconnect,” in *Optical Fiber Communication Conference*, 2019.

[31] Zhang, Wenjia, Ling Ge, Yanci Zhang, Chenyu Liang, and Zuyuan He, “Compressed nonlinear equalizers for 112-gbps optical interconnects: Efficiency and stability,” *Sensors,* vol. 17, p. 4680, 2020.

[32] Zhang, Shaoliang, Fatih Yaman, Kohei Nakamura, Takanori Inoue, Valey Kamalov, Ljupcho Jovanovski, Vijay Vusirikala, Eduardo Mateo, Yoshihisa Inada, and Ting Wang, “Field and lab experimental demonstration of nonlinear impairment compensation using neural networks,” in *Nature communications*, 2019.

[33] Xu, Tianhua, Nikita A. Shevchenko, Yunfan Zhang, Cenqin Jin, Jian Zhao, and Tiegen Liu, “Information rates in Kerr nonlinearity limited optical fiber communication systems,” *Optics Express,* vol. 29, pp. 17428-17439, 2021.

[34] Li, Mingyuan, Wenjia Zhang, Qiao Chen, and Zuyuan He, “High-throughput hardware deployment of pruned neural network based nonlinear equalization for 100-Gbps short-reach optical interconnect,” *Optics Letters,* vol. 46, pp. 4980-4983, 2021.

[35] Renda, Alex, Jonathan Frankle, and Michael Carbin, “Comparing rewinding and fine-tuning in neural network pruning,” in *International Conference on Learning Representations*, 2019.

[36] Weerts, Hilde JP, Andreas C. Mueller, and Joaquin Vanschoren, “Importance of tuning hyperparameters of machine learning algorithms,” in *arXiv preprint arXiv*, 2020.

[37] Pedro J Freire, Michael Anderson, Bernhard Spinnler, Thomas Bex, Jaroslaw E Prilepsky, Tobias A Eriksson, Nelson Costa, Wolfgang Schairer, Michaela Blott, Antonio Napoli, Sergei K Turitsyn, “Towards FPGA Implementation of Neural Network-Based Nonlinearity Mitigation Equalizers in Coherent Optical Transmission Systems,” *European Conference on Optical Communications (ECOC),* 2022.

[38] Pedro J Freire, Antonio Napoli, Diego Arguello Ron, Bernhard Spinnler, Michael Anderson, Wolfgang Schairer, Thomas Bex, Nelson Costa, Sergei K Turitsyn, Jaroslaw E Prilepsky, “Reducing Computational Complexity of Neural Networks in Optical Channel Equalization: From Concepts to Implementation,” *arXiv preprint arXiv:2208.12866,* 2022.

[39] He, Yihui, Ji Lin, Zhijian Liu, Hanrui Wang, Li-Jia Li, and Song Han, “Amc: Automl for model compression and acceleration on mobile devices,” *the European conference on computer vision (ECCV),* pp. 784-800, 2018.

## EC Funding

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 956713