POLITECNICO DI TORINO
Repository ISTITUZIONALE

AI-Driven Detection and Localization of Optical Filtering Soft Failures via Digital Twin

Original
AI-Driven Detection and Localization of Optical Filtering Soft Failures via Digital Twin / Rosso, A., Miotto, E., D'Ingillo, R., Curri, V.. - ELETTRONICO. - (2026), pp. 1-4. (European Conference and Exhibition on Optical Communication 2026 Malaga (Spa) 20–24 September 2026).

Availability:
This version is available at: 11583/3017998 since: 2026-11-16T15:04:20Z

Publisher:
Optica Publishing Group

Published
DOI:

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

29 November 2026

AI-Driven Detection and Localization of Optical Filtering Soft Failures via Digital Twin

Andrea Rosso[(1)], Enrico Miotto[(1)], Rocco D'Ingillo[(1)], Vittorio Curri[(1)]

[(1)] Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Italy, andrea.rosso@polito.it

Abstract *We propose a digital twin-assisted neural network for detecting, localizing, and characterizing soft failures in optical filters. Leveraging optical channel monitor and transceiver metrics, the method achieves high accuracy enabling scalable monitoring in complex optical networks. ©2026 The Authors*

Introduction

The continuous growth of data traffic, driven by cloud computing, immersive services, artificial intelligence (AI), and IoT, is pushing optical networks toward high-density solutions capable of supporting very high data rates with high reliability. To meet these demands, optical signals are dynamically routed using reconfigurable optical add-drop multiplexers (ROADMs) based on wavelength-selective switches (WSSs), typically combined with erbium-doped fiber amplifiers (EDFAs) to compensate for losses.

Under these conditions, optical network digital twins (ONDTs) [1, 2] enable performance prediction through accurate modeling of network elements and are particularly critical in the presence of filtering devices [3, 4], whose impairments must be captured to assess signal performance [5].

As network complexity increases, so does the number of potential failures, especially *soft* failures that degrade performance without disrupting link continuity. Prompt detection and localization are therefore essential to enable timely restoration of the nominal quality of transmission (QoT) [6].

AI has been widely applied in optical networks to predict performance from received signals [7], estimate the optical signal-to-noise ratio (OSNR) using deep neural networks [8, 9], and perform failure prediction and early detection [10, 11]. Recent works have also explored the joint use of AI and ONDTs [12]. Soft failures are well studied in EDFAs [13, 14], while those in optical filters have received limited attention and are often analyzed using an Optical Spectrum Analyzer (OSA) [15].

In this work, we propose a neural network (NN)-based method for failure detection and localization in optical networks affected by filtering impairments, leveraging an ONDT including filtering, transceiver, and equalizer models. To the best of our knowledge, this is the first approach to jointly address detection, localization, and characterization of optical filtering soft failures without relying on OSA measurements, using only a single Optical Channel Monitor (OCM) and SNR derived from BER or Q-factor.

Optical Link Modeling via ONDT

In this work, an ONDT-based software framework is implemented to simulate an optical link, including ASE noise, transceiver, equalization, and filtering effects, using models validated in [5]. The considered line, shown in Fig.1, consists of 10 optical filtering stages (ROADMs) and pre and booster amplifiers introducing ASE noise. This configuration is representative of metro networks, with many filtering elements enabling evaluation of the NN over a wide range of scenarios.

The channel under test (CUT) at 193.5 THz is generated by an optical coherent transceiver whose electrical noise is modeled through an $\mathrm{SNR}_{\mathrm{TRX}}$ parameter, whose value represents the performance of the transceiver for different received optical power (ROP), as in [16]. The CUT roll-off ρ and symbol rate R_s can be tuned based on the desired modulation and spectral occupation. All filters are assumed identical, with the same central frequency as the CUT, and their spectral shapes are defined by two parameters (B_{ch} and BW_{OTF}) according to [17], with B_{ch} restricted to multiples of 6.25 GHz, consistent with the standard WDM grid. Optical amplifiers are assumed to fully compensate for losses and introduce identical ASE noise while operating in transparency, thus contributing equally to the overall $\mathrm{SNR}_{\mathrm{ASE}}$.

The last ROADM (i.e., position 10 in Fig.1) is equipped with an OCM that measures the signal power by integrating the power spectral density (PSD) over 6.25 GHz intervals, as typical for such devices. This yields a set of measurements, each corresponding to the optical power within a 6.25 GHz slot. For data generation, 25 contiguous slots are considered to ensure full signal capture. The OCM values are perturbed by a uniformly distributed error over a 0.1 dB range around the true value to account for measurement uncertainty. In addition to the OCM readings at the last ROADM, the ONDT also provides the *nominal* OCM values, i.e., those under fault-free conditions, as derived from its mathematical models.

In addition to the OCM readings, the transceiver provides the SNR after Fractionally Spaced Equal-

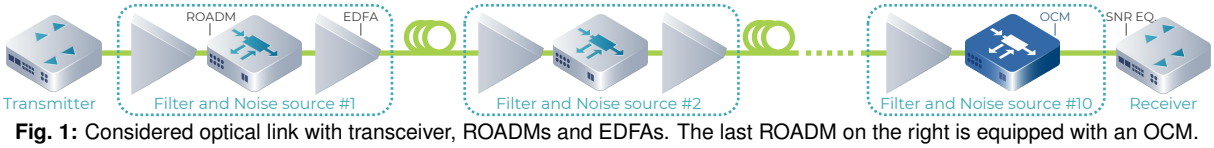

Fig. 1: Considered optical link with transceiver, ROADMs and EDFAs. The last ROADM on the right is equipped with an OCM.

Parameter	Tested values	B_{ch} values [GHz]
R_s [GHz]	31.6 44.0, 50.0 55.0, 63.1 70.0	31.25, 37.5, 43.75 43.75, 50.0, 56.25 56.25, 62.5, 68.75 68.75, 75.0, 81.25
roll-off	0.1, 0.2	–
BW_{OTF} [GHz]	10, 16	–
$\mathrm{SNR}_{\mathrm{ASE}}$ [dB]	10, 12, 14, 16	–
$\mathrm{SNR}_{\mathrm{TRX}}$ [dB]	18, 20, 22	–

Tab. 1: Considered system parameters. For every R_s it is also reported the list of considered B_{ch} values.

ization (FSE) [5]. In real networks, this metric can be derived from BER or Q-factor measurements, enabling straightforward comparison between ONDT predictions and real implementations. FSE offers the implementation of an infinite-length digital equalizer, based on minimum mean square error, and is preferred over finite-length models, which would require additional constraints on the number of taps, beyond the scope of this work.

Dataset construction and NN Architecture

Tab.1 reports the combinations of link parameters selected to generate the dataset used for training, validation, and testing of the deep learning model. By combining all parameter configurations listed in Tab.1, a total of 864 *in-service* conditions is obtained, generating also the corresponding *nominal* OCM readings.

The set of *faulty* conditions is then generated by introducing a soft failure, modeled as a shift of a filter central frequency that causes misalignment with the channel under test, while all other line parameters are kept unchanged. For each case, a new ONDT instance is used to compute the corresponding 25 OCM readings affected by the misalignment. In this work, we assume at most one faulty filter per sample, which simplifies the learning task and reflects typical operational scenarios where faults are rare and isolated. We also assume that the faulty filter shift is an integer multiple of the frequency slot. A failure is simulated by randomly selecting a filter and shifting it by ± 1, ± 2, or ± 3 slots of 6.25 GHz, yielding six possible conditions. For each combination of filtering parameters, assuming at most one faulty filter, the dataset includes 15 randomly sampled soft-failure cases (out of 60 possible single-filter conditions), ensuring diversity while limiting computational cost. Two samples of OCM readings are available in Fig.2. Overall, combining in-service and soft-failure cases results in 13824 unique samples. The dataset is then randomly split into training, validation, and test sets with proportions of 80%, 10%, and 10%, respectively.

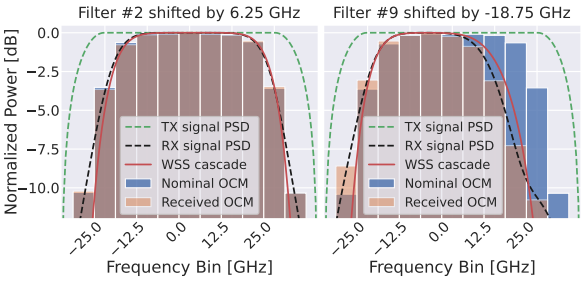

Fig. 2: Examples of OCM readings in case of $R_s = 63.1$ GHz, $B_{\mathrm{ch}} = 68.75$ GHz, $BW_{\mathrm{OTF}} = 16$ GHz, roll-off $\rho = 0.2$, $\mathrm{SNR}_{\mathrm{ASE}} = 16$ dB, $\mathrm{SNR}_{\mathrm{TRX}} = 20$ dB.

For automatic fault detection, a Deep Neural Network (DNN) is adopted, as it can effectively model complex relationships in the data while reducing the need for extensive feature engineering [18]. In the considered setup, the input dimensionality consists of 59 features. These include four parameters related to the signal and filter characteristics (R_s, ρ, B_{ch}, BW_{OTF}), two noise power terms (ASE noise and TRX electrical noise), two SNR values (ASE and TRX), and the post-equalization SNR obtained from the transceiver. In addition, 25 features correspond to the OCM readings, while another 25 represent the difference between the *nominal* OCM readings and the OCM readings provided by the last ROADM. The DNN adopts a multi-task architecture with three output heads: one for fault detection (binary classification), one for fault localization (multi-class), and the last one for shift estimation (regression).

The DNN hyperparameters are optimized using Optuna [19], yielding a configuration with three hidden layers of 512, 1024, and 512 neurons, a batch size of 128 and a learning rate of 4×10^{-4}.

In the considered setup, several accuracy metrics are defined. Given a dataset of size N, and N_f faulty samples, the fault accuracy is defined as the ratio of correct fault/no-fault predictions to N, while position and shift accuracies are computed over faulty samples as the fraction of correctly identified positions and shifts over N_f, respectively. The end-to-end accuracy is evaluated over N, counting a prediction as correct if it identifies either a no-fault condition or, in the presence of a fault, both its position and shift. Three loss functions are used: binary cross-entropy for fault detection, categorical cross-entropy for fault position, and mean squared error for shift estimation. The overall loss is a weighted average with weights optimized using Optuna. Its evolution, shown in Fig.3, indicates stable convergence, with decreasing training and validation loss and steadily improving accuracy. The close alignment of the training and validation loss curves suggests good generalization.

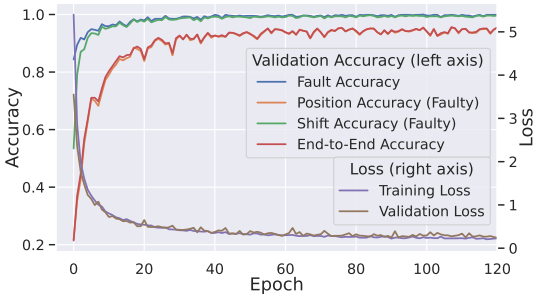

Fig. 3: Validation accuracy and loss progress during training.

fault/no fault	position	shift	e2e
0.9954	0.9507	0.9914	0.9502

Tab. 2: Accuracy results on the test set.

	precision	recall	F1 score
macro	0.9726	0.9981	0.9850
micro	0.9964	0.9964	0.9964

Tab. 3: Results for fault detection on the test set.

Results

This section presents the results obtained on the test set after training and validation of the NN. Accuracy metrics are reported in Tab.2, while fault detection results are shown in Tab.3. The accuracy exceeds 0.99 for fault detection and shift evaluation, and 0.95 for fault localization. This reflects the higher complexity of localization compared to detection and shift estimation. Notably, the high shift accuracy indicates that once a fault is detected, its magnitude can be reliably estimated. Consequently, the end-to-end accuracy is close to the position accuracy. Due to the strong class imbalance in the case of fault detection, with only 864 no-fault cases out of 13,824 total samples, accuracy alone is not sufficient to fully assess the classifier performance. For this reason, macro and micro averaged precision, recall, and F1-score are also considered. As reported in Tab.3, all these metrics are higher than 0.97, confirming the robustness of the model.

The results on accuracy demonstrate the effectiveness of the proposed approach. To further assess the strengths and limitations of the model, we analyze the multi-class confusion matrix (Fig.4) for fault localization. The matrix shows strong classification performance, with most samples concentrated along the main diagonal. The No Fault class is perfectly recognized, while fault-location classes achieve high accuracy overall. Misclassifications are limited and mainly occur between adjacent positions, suggesting that errors arise between nearby fault locations with similar features rather than distant ones. Overall, the model achieves high accuracy in both fault detection and localization, with only minor confusion among neighboring classes towards the transmitter side.

From the first column of the confusion matrix, some faulty cases are misclassified as No Fault, corresponding to false negatives in the detection task. Since the matrix is normalized per class, we

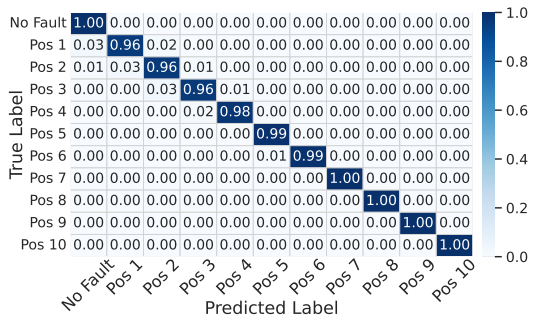

Fig. 4: Normalized Confusion Matrix for failure localization.

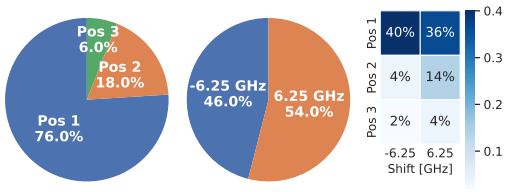

Fig. 5: Analysis of false negative samples.

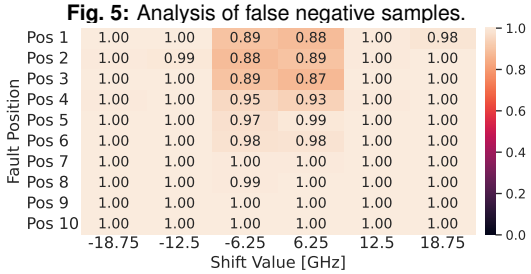

Fig. 6: Accuracy on faulty samples by position and shifts.

analyze these cases to assess the impact of fault position and shift on performance. The results, shown in Fig.5, indicate that the closer the faulty filter is to the transmitter, the higher the probability of a false negative classification. This behavior can be explained by the cascade of filtering stages: distortions introduced early in the link are partially reshaped by subsequent filters, reducing their visibility at the receiver. In contrast, faults closer to the receiver directly affect the observed spectrum. These false negatives also propagate to the localization and shift estimation tasks, as these metrics are computed over samples with ground-truth faults, making the evaluation more representative of end-to-end performance.

After analyzing false negatives, we examine faulty samples to assess the impact of frequency shift on classification. This analysis is summarized in the heatmap in Fig.6. As expected, the most critical cases occur for shifts of ± 6.25 GHz, which are comparable to the OCM frequency resolution, making them difficult to distinguish from nominal conditions. This increases the likelihood of misclassifications or missed detection. As already considered for the confusion matrix and the pie charts, performance also depends on the fault position, with early-stage faults showing slightly lower detection and localization accuracy.

Conclusions

In this work, we showed that the combination of an ONDT and a multi-task DNN enables accurate fault detection and localization in complex network configurations, using only a single OCM and transceiver-based measurements. Simulation results indicate that performance depends on the failure position, while maintaining high accuracy across most scenarios. Future work will focus on experimental validation and on extending the proposed approach to more complex scenarios, including multiple simultaneous failures. Overall, the proposed approach demonstrates its strong potential for practical deployment in monitoring and fault management of optical networks.

Acknowledgments

This publication has received funding from the EU Horizon project SENSEI (GA: 101189545). The authors wish to thank the Telecom Infra Project.

References

[1] V. Curri, "GNPy model of the physical layer for open and disaggregated optical networking [Invited]", *Journal of Optical Communications and Networking*, Jun. 2022. DOI: 10.1364/JOCN.452868.

[2] H. Nishizawa, T. Mano, K. Anazawa, T. Matsumura, T. Sasai, M. Namiki, D. Briantcev, R. Ambrosone, E. L. Rouzic, S. Melin, O. Gonzalez-de-Dios, J. P. Fernandez-Palacios, X. Zhang, K. Akahoshi, G. Grammel, A. D'Amico, G. Borraccini, M. Ruffini, D. Kilper, and V. Curri, *Optical Network Digital Twin – Commercialization Barriers, Value Proposition, Early Use Cases, and Challenges*, arXiv:2511.06368 [eess], Nov. 2025. DOI: 10.48550/arXiv.2511.06368.

[3] R. Ayassi, X. Yang, C. Sun, Y. Pointurier, and G. Charlet, "Filtering Impairments-Aware Digital Twin for SNR Prediction over Network Life-Cycle", *Optical Fiber Communications Conference and Exhibition (OFC)*, Apr. 2025. DOI: 10.1364/OFC.2025.W4G.4.

[4] E. Miotto, A. Rosso, E. Virgillito, S. Straullu, A. Castoldi, A. Bovio, F. M. Rodriguez, and V. Curri, "QoT Impairments Induced by Statistical Filtering Variations with a Realistic Equalizer", in *2025 European Conference on Optical Communications (ECOC)*, Copenhagen, Denmark: IEEE, Sep. 2025, pp. 1–4, ISBN: 979-8-3315-9531-9. DOI: 10.1109/ECOC66593.2025.11263006.

[5] E. Miotto, A. Rosso, S. Straullu, E. Virgillito, R. Corsini, A. Bovio, and V. Curri, *Generalized Analytical Modeling of Filtering Penalties and Equalization in Coherent Optical Links*, Mar. 2026. DOI: 10.36227/techrxiv.177273899.97147441/v1.

[6] I. Sartzetakis, K. K. Christodoulopoulos, and E. M. Varvarigos, "On reducing optical monitoring uncertainties and localizing soft failures", in *2017 IEEE International Conference on Communications (ICC)*, Paris, France: IEEE, May 2017, pp. 1–6, ISBN: 978-1-4673-8999-0. DOI: 10.1109/ICC.2017.7997218.

[7] Xiaoxia Wu, J. Jargon, R. Skoog, L. Paraschis, and A. Willner, "Applications of Artificial Neural Networks in Optical Performance Monitoring", *Journal of Lightwave Technology*, vol. 27, no. 16, pp. 3580–3589, Aug. 2009, ISSN: 0733-8724, 1558-2213. DOI: 10.1109/JLT.2009.2024435.

[8] T. Tanimura, T. Hoshida, J. C. Rasmussen, M. Suzuki, and H. Morikawa, "OSNR monitoring by deep neural networks trained with asynchronously sampled data", Niigata, Japan: IEEE, Jul. 2016.

[9] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An Overview on Application of Machine Learning Techniques in Optical Networks", *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1383–1408, 2019, ISSN: 1553-877X, 2373-745X. DOI: 10.1109/COMST.2018.2880039.

[10] S. Cruzes, "Failure Management Overview in Optical Networks", *IEEE Access*, vol. 12, pp. 169 170–169 193, 2024, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2024.3498704.

[11] F. Musumeci, C. Rottondi, G. Corani, S. Shahkarami, F. Cugini, and M. Tornatore, "A Tutorial on Machine Learning for Failure Management in Optical Networks", *Journal of Lightwave Technology*, vol. 37, no. 16, pp. 4125–4139, Aug. 2019, ISSN: 0733-8724, 1558-2213. DOI: 10.1109/JLT.2019.2922586.

[12] M. C. Mohamed, R. Ambrosone, M. U. Masood, G. Malik, S. Straullu, A. Nespola, S. K. Bhyri, A. Napoli, J. Pedro, G. M. Galimberti, W. Wakim, and V. Curri, "Failure Prediction in Optical Transport Networks Through the Integration of Digital Twins and Deep Learning", *Journal of Lightwave Technology*, vol. 44, no. 8, pp. 2964–2973, Apr. 2026, ISSN: 0733-8724, 1558-2213. DOI: 10.1109/JLT.2026.3655186.

[13] M. C. Mohamed, M. U. Masood, I. C. Dipto, R. Ambrosone, G. Malik, S. Straullu, S. K. Bhyri, G. M. Galimberti, J. Pedro, A. Napoli, W. Wakim, and V. Curri, "Network Data Based Transfer Learning Failure Prediction Agent Pre-Trained Using Digital Twin", in *2025 International Conference on Optical Network Design and Modeling (ONDM)*, Pisa, Italy: IEEE, May 2025, pp. 1–6, ISBN: 978-3-903176-67-6. DOI: 10.23919/ONDM65745.2025.11029222.

[14] M. C. Mohamed, R. Ambrosone, R. D'Ingillo, M. U. Masood, G. Malik, S. Straullu, S. K. Bhyri, J. Pedro, A. Napoli, G. M. Galimberti, W. Wakim, and V. Curri, "Leveraging Digital Twins to Build Comprehensive Network Failure Databases for Predictive Machine Learning in Optical Transport Networks", in *2025 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, Delhi, India: IEEE, Dec. 2025, pp. 1–6, ISBN: 979-8-3315-2681-8. DOI: 10.1109/ANTS66931.2025.11430070.

[15] L. Velasco, B. Shariati, A. P. Vela, J. Comellas, and M. Ruiz, "Learning from the Optical Spectrum: Soft-Failure Identification and Localization [Invited]", in *Optical Fiber Communication Conference*, San Diego, California: OSA, 2018, W1G.1, ISBN: 978-1-943580-38-5. DOI: 10.1364/OFC.2018.W1G.1.

[16] T. Mano, Y.-K. Huang, G. Borraccini, E. Ip, A. D'Amico, Z. Wang, H. Nishizawa, G. Zussman, T. Chen, T. Wang, K. Asahi, D. Kilper, V. Curri, and K. Takasugi, "Modeling the Input Power Dependency of Transceiver BER-ONSR for QoT Estimation", in *Optical Fiber Communication Conference (OFC) 2024*, San Diego California: Optica Publishing Group, 2024, M1H.4, ISBN: 978-1-957171-32-6. DOI: 10.1364/OFC.2024.M1H.4.

[17] C. Pulikkaseril, L. A. Stewart, M. A. F. Roelens, G. W. Baxter, S. Poole, and S. Frisken, "Spectral modeling of channel band shapes in wavelength selective switches", *Optics Express*, vol. 19, no. 9, p. 8458, Apr. 2011, ISSN: 1094-4087. DOI: 10.1364/OE.19.008458.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature14539.

[19] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A Next-generation Hyperparameter Optimization Framework", in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage AK USA: ACM, Jul. 2019, pp. 2623–2631, ISBN: 978-1-4503-6201-6. DOI: 10.1145/3292500.3330701.