Machine Learning Agents Leveraging Digital Twins for Failure Prediction in Optical Networks

Machine Learning Agents Leveraging Digital Twins for Failure Prediction in Optical Networks

Mashboob Cheruvakkadu Mohamed
Politecnico di Torino, Italy
mashboob.cheruvakkadu@polito.it

Muhammad Umar Masood
Politecnico di Torino, Italy
muhammad.masood@polito.it

Renato Ambrosone
Politecnico di Torino, Italy
renato.ambrosone@polito.it

Gulmina Malik
Politecnico di Torino, Italy
gulmina.malik@polito.it

Rocco D'Ingillo
Politecnico di Torino, Italy
rocco.dingillo@polito.it

Stefano Straullu
Links Foundation, Torino, Italy
stefano.straullu@linksfoundation.com

Sai Kishore Bhyri
Infinera India
sbhyri@infinera.com

Gabriele Maria Galimberti
Infinera USA
ggalimberti@infinera.com

João Pedro
Infinera Unipessoal Lda, Portugal
jpedro@infinera.com

Antonio Napoli
Infinera Germany
anapoli@infinera.com

Walid Wakim
Infinera USA
wwakim@infinera.com

Vittorio Curri
Politecnico di Torino, Italy
vittorio.curri@polito.it

***Abstract*—This study proposes an advanced framework for predicting optical amplifier failures in optical networks by integrating Digital Twins (DT) and Machine Learning (ML). Utilizing the GNPy open-source framework, DTs replicate amplifier behavior under various conditions, resembling faults and captures the network conditions and performance metrics of the optical networks. The telemetry data generated from these simulations represents both short-term dynamics and long-term trends in amplifier performance, enabling the training of a Long Short-Term Memory (LSTM) model. The ML model demonstrates an 'amplifier fault level' classification accuracy of 98 %, effectively identifying soft failures and assessing fault severity. By leveraging the ability to model complex fault scenarios in a controlled environment, the framework provides a comprehensive solution for generating datasets that are otherwise difficult to obtain from live networks. This approach enables early detection and intervention, minimizes service disruptions, and enhances network reliability. The integration of DTs and LSTM-based ML offers a scalable and data-driven solution for improving the resilience, efficiency, and operational continuity of modern optical communication systems.**

***Index Terms*—Machine Learning, Digital Twin, Optical Amplifiers Fault Detection, Soft failures, Proactive Failure Management, Network Resilience.**

I. Introduction

Today's optical networks are constrained by rigid operational approaches, which can restrict network availability and efficiency, leading to a longer time to restore interruptions in their network infrastructure [1]. Long network down-time can cause heavy monetary loss for the operators and provides the motivation to invest in proactive maintenance strategies for optical networks. Effective prediction of failures can result in the avoidance of service disruptions, thus enhancing the availability and stability of these networks. With early detection of soft failures, the desired optical signal-to-noise ratio (OSNR), bit error rate (BER), and high Quality of Transmission (QoT) can be achieved while ensuring the continuity of service in the optical network. Effective failure management – by detecting and localizing soft failures before they lead to hard failures – can resort to various ML techniques [2].

ML provides powerful methods for automating the detection and identification of soft failures, greatly improving failure management and improving network reliability [3]. For the best performance of the ML model, it is crucial that the model is trained with a proper dataset. The collection of failure datasets for the training of ML from live networks is limited by the fact that the occurrence of failure events is irregular in real networks, and may also only contemplate a sub-set of the possible failure events. This can lead to imbalance datasets raising concerns about the reliability and effectiveness in the implementation of these ML models for practical use cases.

Recent research highlights the concern about the impact of unbalanced data sets, due to the scarcity of certain failure types, in the trained ML model [4], [5]. A potential alternative for the generation of training data is the use of Digital Twins (DT). In [6], telemetry twin based on soft failure localization is demonstrated in which the faulty network device is identified from the instantaneous values from the telemetry without addressing the time-series properties. This limits any possibility for the differentiation between the instant spike due to network dynamics or the persistence of the real fault in the network device. Thus, it can lead to compromised usability, given the extent to which real-time synchronization of the telemetry twin with the actual network is achievable.

The use of historical data from the DT is proposed in [7]. However, the collection of historical data from the DT of a real network depends on the occurrence of failures in networks, which is completely uneven and infrequent, prolonging the time required to accumulate all types of failure data, rendering this solution impractical at present. This study presents a

This publication has received funding and support from the project PNRR-NGEU (MUR–DM117/2023) and from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101092766 (ALLEGRO Project).

979-8-3315-2042-7/25/$31.00 ©2025 IEEE

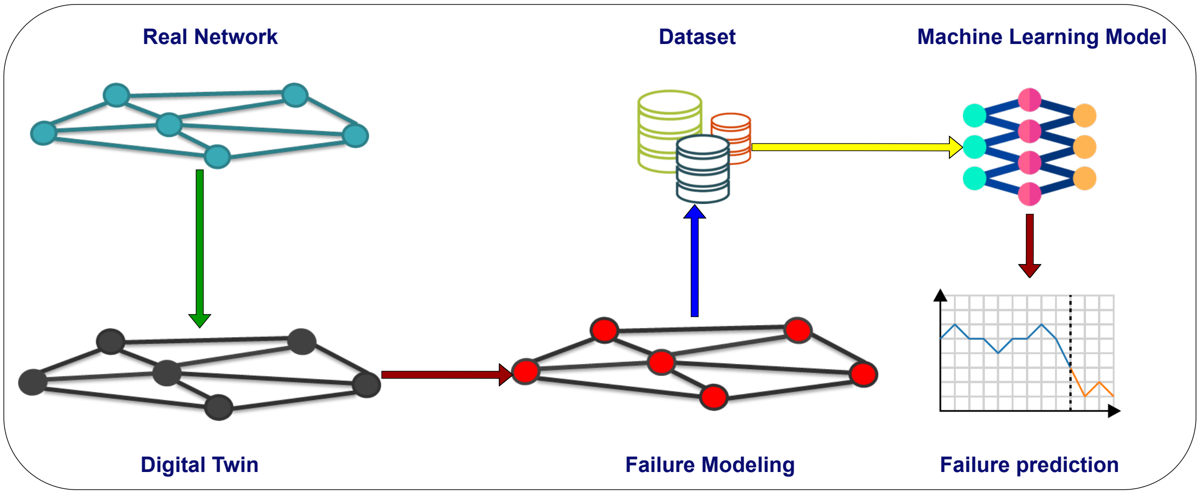

Fig. 1: Fault forecasting using failure modeled copy of DT (GNPy), a schematic representation.

novel method for creating a comprehensive failure dataset from QoT DT, GNPy. It involves modeling various types of failure to assess their impact on transmission metrics such as Generalized Signal to Noise Ratio (GSNR), OSNR, BER, and channel power.

The framework of the overall process of dataset generation using the DT and ML model for failure prediction is depicted in Fig. 1. Several of the existing research works on ML for soft-failures demonstrate high accuracy using complex models trained on partial / limited data. This limitation may translate into a compromised ability to identify and localize different types of failures and their evolution over time [4]. In this article, we demonstrate ML technique (LSTM) addressing the time-series properties of failure scenarios by leveraging a comprehensive dataset generated by simulating the evolution of failures using GNPy. The following outlines the structure of the remaining sections of this paper. DT and data acquisition is demonstrated in Section II. In Section III, the ML framework is explained. Furthermore, performance evaluation of the model is elaborated in Section IV. Finally, concluding remarks are presented in Section V.

II. Digital Twin and Data Acquisition

A. GNPy Powered Digital Twin of Optical Multiplexed Section

The concept of DT holds the potential to transform optical communication networks by facilitating comprehensive lifecycle management, such as planning, optimization, forecasting, etc. There is significant advancement in DT applications for optical networks, spanning from conceptualization, design, simulation, and experimental validation. Contributing as a virtual counterpart in real time to physical optical networks, the DT is divided into three key components [8]:

1) **Digital Model:** A virtual representation of the physical system, replicating its characteristics, parameters, and behavior. In optical networks, this includes elements such as amplifiers, fibers, and their associated parameters.
2) **Real-Time Data Connection:** A bidirectional flow of data between the physical system and its digital counterpart, enabling the DT to reflect the current state of the system. These data can include telemetry, sensor readings, or live measurements.
3) **System Control and Optimization:** The ability to actively manage and optimize the physical system through the DT, leveraging real-time insights and simulations. This includes adjusting operational parameters, mitigating failures, or improving overall performance.

The GNPy framework, which adopts the Gaussian Noise (GN) model, focuses on QoT estimation (QoT-E) in optical networks using coherent transceivers. GNPy models transparent lightpaths as channels with additive white GN, making it suitable as a vendor-neutral DT for open network design and management [9]. GNPy, as an open-source application, is being widely used to estimate the capacity and performance of a deployed network, ensuring more efficient network design and operation [10].

The network topology, which includes components such as transponders, Reconfigurable Optical Add/Drop Multiplexers (ROADMs), amplifiers, and optical fibers, is modeled with the DT (GNPy), as illustrated in Fig. 2. In this work, a 520 km link of single-mode fiber comprising a total of eight spans of 65 km each is considered. The simulated network includes the virtual twin of commercially deployed amplifiers consisting of seven In-Line Amplifiers (ILA), booster and pre-amplifier in both directions.

B. Amplifier Failure: Data Acquisition

Optical amplifiers play a crucial role in long-distance optical transmission by directly amplifying optical signals without the need for optical-to-electrical conversion. Performance degradation in optical amplifiers can significantly impact OSNR

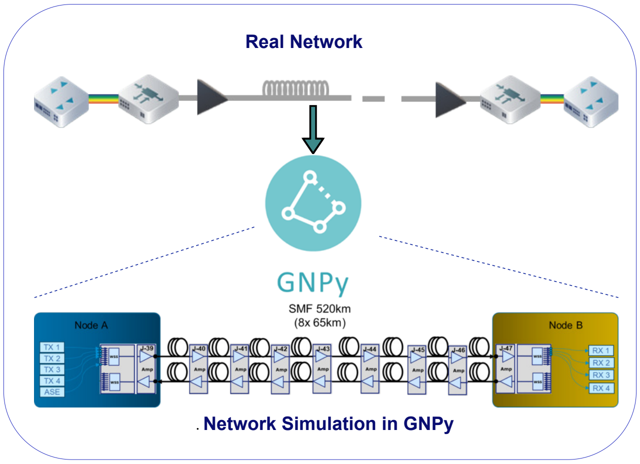

Fig. 2: Network topology including device characteristics, fiber span length etc. simulated in GNPy, an illustration.

leading to low QoT. The common issues that can arise with optical amplifiers and lead to failures that negatively affect network performance are as follows [1]:

- **Amplifier Gain Reduction:** Gain degradation in amplifier over time can be caused by aging pump lasers, saturation of gain media, resulting in reduced overall gain, and lower efficiency of the amplifier.
- **Noise Increment:** The Amplified Spontaneous Emission (ASE) noise can increase due to instabilities, inefficiencies, aging, or wavelength drift of the pump laser.
- **Output Power Anomaly:** Fluctuations in output power such as spikes or drops can be caused by variation in pump current or pump failures.

The continued presence of these issues can result in critical failures that disrupt services. Thus, it is essential to continuously monitor amplifier failures and implement proactive failure management strategies, as amplifiers are vital components in an Optical Line System (OLS) due to their role in maintaining the required performance-level.

The network topology is simulated with GNPy to estimate the effects of physical layer impairments such as ASE noise, non-linear interference (NLI) and accumulated dispersion, on signal quality. The tool accepts as input a description of the network topology, which includes links, nodes, and optical components, as well as equipment parameters like amplifier gain, noise figure, and transponder specifications. It utilizes this information to calculate metrics such as GSNR, OSNR, and other QoT parameters.

GNPy performs signal propagation calculations by simulating the interactions of optical signals with network elements, including fibers, amplifiers, and ROADMs. It accounts for physical layer characteristics like fiber attenuation, span lengths, wavelength-dependent gain profiles, and nonlinear distortions. By iteratively applying the GN model over the network path, GNPy emulates the cumulative effects of impairments on the transmitted signal.

We utilized GNPy to generate the training dataset by systematically varying the input network topology and associated parameters across multiple simulations. This approach allowed us to simulate different network configurations and operating conditions, including fault scenarios, which are challenging to capture from real-world deployments due to their rarity and the complexity of data collection in live systems.

By modifying elements such as amplifier gains, transponder settings, and fault conditions, GNPy enabled the emulation of a wide range of network behaviors and signal quality metrics. Each simulation produced detailed QoT data, including GSNR and OSNR values, which were added to the dataset. Subsequently, slow gain degradation is induced in the amplifier by varying the gain in small incremental steps and adjusting the Variable Optical Attenuator (VOA) in parallel to compensate for the power loss so that it represents silent failures in the network which are not detectable directly by the power monitors. Three fault levels are assigned according to the degradation value, until 2 dB is considered as no fault, between 2 dB and 3 dB as minor fault, and above 3 dB as major fault. To avoid service interruptions, the maximum fault of an amplifier must remain within the VOA limit, as it aims to predict soft failures. Exceeding this threshold may lead to hard failures.

The change in transmission parameters due to the impact of failures is assessed. The step sizes are chosen randomly, ensuring that the overall deterioration accumulates gradually to simulate field defects. The particular step size for different time periods is set randomly, in order to simulate the network variations. The time-series dataset used for training of the ML model comprise of variations in GSNR that correspond to the degradation values of individual amplifiers, each having different gain and tilt values.

The generated dataset serves as a valuable resource for training and validating ML models designed to predict and diagnose network performance, optimize resource allocation, and enhance fault localization and mitigation strategies. However, the synthetic dataset may not represent exactly the real-world condition. Using GNPy's flexibility, it is ensured to maximum possible extent that the dataset includes realistic yet diverse scenarios to improve the robustness and generalizability of the ML models.

III. Machine Learning Framework Overview

In this section, the background on the LSTM ML technique is outlined, detailing its application to the amplifier fault prediction and classification of fault levels.

A. *LSTM - Background and Relevance*

LSTM is a special type of Recurrent Neural Network (RNN) specifically formulated to tackle the problem of the vanishing gradient. The problem of vanishing gradient arises while training if the loss function gradients pertaining to the model parameters diminishes during backward propagation. Very small gradients can be a challenging problem in updating the weights of the model, which can be addressed by

LSTM [11], [12]. LSTM networks were originally proposed by Hochreiter and Schmidhuber [13]. Leveraging LSTM in ML significantly boosts forecasting accuracy by incorporating both recent and historical time-series data. [14]. LSTM is a popular and effective model for input features with sequential dependencies [11].

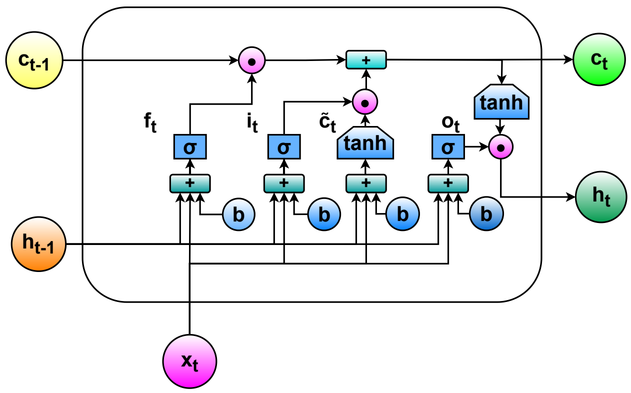

Fig. 3: LSTM basic structure; x_t: input vector, c_{t-1}: memory from the previous block, h_{t-1}: output of the previous block, c_t: memory from the current block, h_t: output of the current block, σ: sigmoid, $\tanh$: hyperbolic tangent, b: bias, $\cdot$: element-wise multiplication, $+$: element-wise summation.

The central component of the LSTM technique is the memory cell, consisting of the input, output, forget gates, and the activation function [13]. The flow of information to memory cells is managed by these gates keeping past data over time, and forecasts are given based on interactions between input values, long-term and short-term memories [1]. Fig. 3 shows the basic structure of LSTM. The core of the model is a bidirectional LSTM, which processes the sequence data in both the forward and backward directions. A standard LSTM model computes its output at each time step as follows:

- Forget Gate:
$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$
- Input Gate:
$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$
- Cell State Update:
$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$
- Cell State:
$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t$$
- Output Gate:
$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$
- Hidden State:
$$h_t = o_t \cdot \tanh(c_t)$$

where:

- W_f, W_i, W_c, W_o: Weight matrices for the forget gate, input gate, cell state, and output gate, respectively, with respect to the input x_t
- U_f, U_i, U_c, U_o: Weight matrices for the forget gate, input gate, cell state, and output gate, respectively, with respect to the hidden state h_{t-1}
- b_f, b_i, b_c, b_o: Bias terms for the forget gate, input gate, candidate cell state, and output gate, respectively.

In a bidirectional LSTM, the input sequence is processed in two ways, from past to future (forward LSTM) and from future to past (backward LSTM). This results in a concatenated output of both hidden states:

$$\hat{h}_t = [h_t^{\text{forward}}, h_t^{\text{backward}}]$$

where h_t^{forward} and h_t^{backward} are the hidden states from the forward and backward LSTM layers.

B. Amplifier Failure Prediction and Fault Estimation

The generated fault data pertaining to the amplifier failure, as described in Section II.B, provides the dataset required for the training of the ML model. The dataset is divided into 80% training data and 20% test data. The training is carried out through 100 epochs. The ML framework is designed for hierarchical multi-class classification using sequential data. The model predicts the faulty amplifier and further classifies the degradation level of the faulty amplifier. It employs a Bidirectional LSTM network to process temporal dependencies in sequences, combined with dense layers for classification. The LSTM layers are followed by dropout layers to regularize the network and prevent overfitting. The model is optimized using the Adam optimizer [15], which adapts the learning rate based on the gradients of the loss function with respect to the model parameters. During training, early stopping is used to prevent overfitting by terminating the training process if the validation loss stops improving for a specified number of epochs. The learning rate scheduling is also used to adjust the learning rate during training. After training, the model performance is evaluated using standard classification metrics. The confusion matrix is generated showing the counts of true positives, false positives, true negatives, and false negatives for each class. This confusion matrix can assess the model performance across each individual class. The classification report, which provides precision [11], recall [11], and F1-score [11] is compiled for each class, giving more detailed insight into the model performance on different types of errors.

IV. Machine Learning Model Performance Evaluation

Effective fault detection and classification are essential to ensure the reliability and performance of optical network in operation. As described, this study employs LSTM to accurately identify faulty amplifiers and also categorize the gradual gain degradation into distinct fault levels. Fig. 4

illustrates the progression of training of the proposed ML model, showing the trends in accuracy (solid blue line) and loss (dashed red line) over 100 epochs. The figure shows

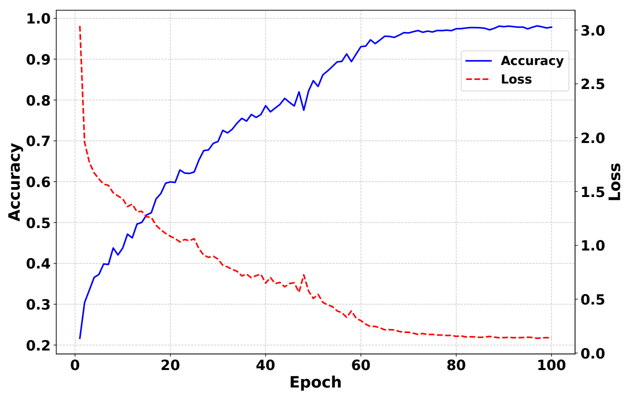

Fig. 4: Model Accuracy and Loss Over Training Epochs.

the relation of accuracy vs. loss over the training of the ML model. It demonstrates a clear trend where accuracy steadily increases while loss consistently decreases, reflecting the model's effective learning process. In the early epochs, rapid improvements in accuracy and reduction in loss are observed, indicating efficient parameter optimization in the initial training process. After around 70 epochs, the model tends to stabilize with a continuous slight improvement in the accuracy and a reduction in the loss. The accuracy vs. loss trend confirm the model's convergence, enabling accurate amplifier fault detection and the subclassification of the fault level in each amplifier.

The evaluation results, including classification performance metrics and confusion matrix are depicted in Fig. 5 and Fig. 6 respectively. Each amplifier is identified by a unique identifier name, such as ila$x$, where x shows the number of ILA. The ML model first detects the faulty amplifier and subsequently classifies the gain degradation into three distinct fault levels: 0, 1, and 2, as detailed in Table I. Labels are used to include both the amplifier identifier and the fault level; for instance, ilax_y refers to the x amplifier with fault level y.

TABLE I: Amplifier gain degradation

Fault level	Gain degradation(dB)	Severity
0	< 2	No fault
1	2-3	Minor fault
2	> 3	Major fault

The ML model achieved an overall classification accuracy of 98%, indicating a strong generalization in the dataset. The classification performance of the ML model was evaluated

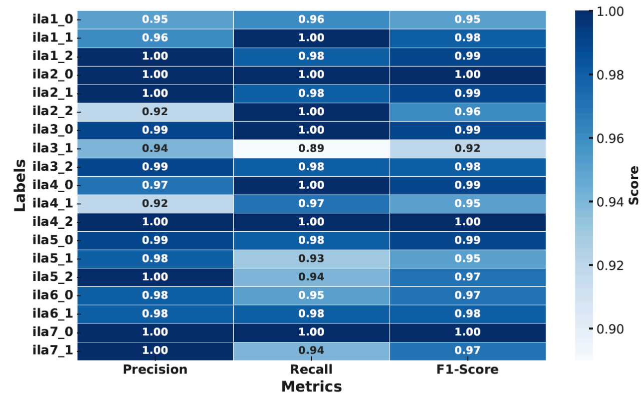

Fig. 5: Classification performance metrics across fault levels for ILAs.

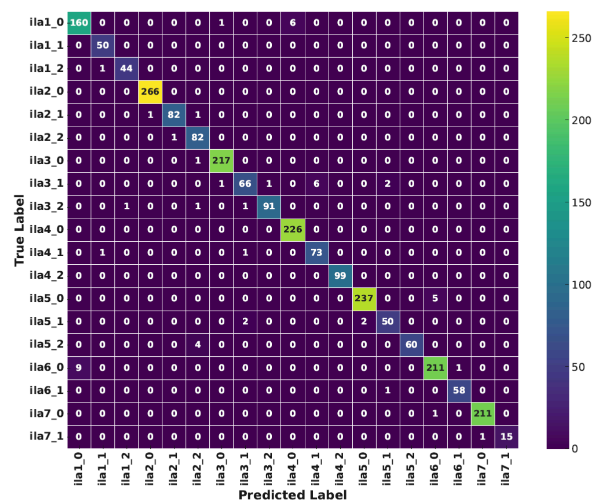

Fig. 6: Confusion matrix showing model performance in predicting amplifier fault levels.

to assess its effectiveness in detecting and classifying fault levels in optical amplifiers. The evaluation metrics include precision, recall, and F1-scores across most classes, which are critical indicators of the model's ability to generalize and provide accurate predictions across all fault levels. In Fig. 5, the model demonstrated consistently high performance, with precision, recall, and F1-scores near or exceeding 0.90 for most classes. Specifically, the model achieved perfect classification performance (precision, recall, and F1-score of 1.00) for labels such as ila2_0, ila4_2, and ila7_0, highlighting its ability to correctly classify these fault levels. For ila3_1, the model showed slightly lower recall, yet still maintained strong performance with precision and F1-scores of 0.94 and 0.92 respectively, demonstrating the model's robustness. The macro average across all classes was 0.97 for precision, recall, and F1-score, suggesting that the model performs well across all fault levels without favoring any particular class. Furthermore, the weighted averages for precision, recall, and F1-score were 0.98, reflecting the effectiveness of the model. Overall, the model demonstrates excellent performance, achieving both high classification accuracy and balanced predictive power across a diverse set of fault levels.

V. Conclusion

In summary, this paper introduced an innovative method for predicting amplifier failures in optical networks through the use of DT and machine learning technique. By constructing DTs that mirror real-time network conditions, the proposed approach enables the early prediction of amplifier failures. This study focused on the application of LSTM, which is trained on DT-time series data of the amplifier faults to predict both short- and long-term fluctuations. The model has been shown to achieve prediction accuracy of 98%, demonstrating its ability to forecast failures with high precision. By anticipating potential failures before they occur, this methodology provides the possibility of a proactive approach to managing and minimizing network disruptions, ultimately improving the reliability of optical networks.

References

[1] S. Cruzes, "Failure management overview in optical networks," *IEEE Access*, vol. 12, pp. 169 170–169 193, 2024.

[2] F. Musumeci, C. Rottondi, G. Corani, S. Shahkarami, F. Cugini, and M. Tornatore, "A tutorial on machine learning for failure management in optical networks," *J. Lightwave Technol.*, vol. 37, no. 16, pp. 4125–4139, Aug 2019. [Online]. Available: https://opg.optica.org/jlt/abstract.cfm?URI=jlt-37-16-4125

[3] D. Wang, Y. Wang, X. Jiang, Y. Zhang, Y. Pang, and M. Zhang, "When large language models meet optical networks: Paving the way for automation," *Electronics*, vol. 13, no. 13, 2024. [Online]. Available: https://www.mdpi.com/2079-9292/13/13/2529

[4] L. Z. Khan, J. ao Pedro, N. Costa, A. Sgambelluri, A. Napoli, and N. Sambo, "Model and data-centric machine learning algorithms to address data scarcity for failure identification," *J. Opt. Commun. Netw.*, vol. 16, no. 3, pp. 369–381, Mar 2024. [Online]. Available: https://opg.optica.org/jocn/abstract.cfm?URI=jocn-16-3-369

[5] C. Xing, C. Zhang, B. Ye, D. Wang, Y. Jia, J. Li, and M. Zhang, "Failure data augmentation for optical network equipment using time-series generative adversarial networks," in *Optical Fiber Communication Conference (OFC) 2023*. Optica Publishing Group, 2023, p. M3G.4. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=OFC-2023-M3G.4

[6] K. S. Mayer, J. A. Soares, R. P. Pinto, C. E. Rothenberg, D. S. Arantes, and D. A. A. Mello, "Machine-learning-based soft-failure localization with partial software-defined networking telemetry," *J. Opt. Commun. Netw.*, vol. 13, no. 10, pp. E122–E131, Oct 2021. [Online]. Available: https://opg.optica.org/jocn/abstract.cfm?URI=jocn-13-10-E122

[7] Q. Zhuge, X. Liu, Y. Zhang, M. Cai, Y. Liu, Q. Qiu, X. Zhong, J. Wu, R. Gao, L. Yi, and W. Hu, "Building a digital twin for intelligent optical networks," *J. Opt. Commun. Netw.*, vol. 15, no. 8, pp. C242–C262, Aug 2023. [Online]. Available: https://opg.optica.org/jocn/abstract.cfm?URI=jocn-15-8-C242

[8] D. Wang, Y. Song, Y. Zhang, X. Jiang, J. Dong, F. N. Khan, T. Sasai, S. Huang, A. P. T. Lau, M. Tornatore, and M. Zhang, "Digital twin of optical networks: A review of recent advances and future trends," *Journal of Lightwave Technology*, vol. 42, no. 12, pp. 4233–4259, 2024.

[9] V. Curri, "Gnpy model of the physical layer for open and disaggregated optical networking [invited]," *Journal of Optical Communications and Networking*, vol. 14, no. 6, pp. C92–C104, 2022.

[10] A. Ferrari, M. Filer, K. Balasubramanian, Y. Yin, E. Le Rouzic, J. Kundrat, G. Grammel, G. Galimberti, and V. Curri, "Gnpy: an open source application for physical layer aware open optical networks," *Journal of Optical Communications and Networking*, vol. 12, no. 6, pp. C31–C40, 2020.

[11] L. E. Kruse, S. Kühl, A. Dochhan, and S. Pachnicke, "Experimental validation of machine learning-based joint failure management and quality of transmission estimation," *IEEE Photonics Journal*, vol. 15, no. 6, pp. 1–9, 2023.

[12] Q. Wang, R.-Q. Peng, J.-Q. Wang, Z. Li, and H.-B. Qu, "Newlstm: An optimized long short-term memory language model for sequence prediction," *IEEE Access*, vol. 8, pp. 65 395–65 401, 2020.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 11 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[14] S. Giaremis, N. Nader, C. Dawson, C. Kaiser, E. Nikidis, and H. Kaiser, "Storm surge modeling in the ai era: Using lstm-based machine learning for enhancing forecasting accuracy," *Coastal Engineering*, vol. 191, p. 104532, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378383924000802

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017. [Online]. Available: https://arxiv.org/abs/1412.6980